

10-667484

```
;      OPERATING SYSTEM: <Unknown>
;      SOFTWARE: ASCII
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/107,532A
;      FILING DATE: 30-Jun-1998
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 60/085,598
;      FILING DATE: 14 May 1998
;      APPLICATION NUMBER: 60/051571
;      FILING DATE: July 2, 1997
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Ariniello, Pamela Deneke
;      REGISTRATION NUMBER: 40,489
;      REFERENCE/DOCKET NUMBER: GTC-012
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (781)893-5007
;      TELEFAX: (781)893-8277
;      INFORMATION FOR SEQ ID NO: 4677:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 358 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      HYPOTHETICAL: YES
;      ORIGINAL SOURCE:
;      ORGANISM: Enterococcus faecium
;      FEATURE:
;      NAME/KEY: misc_feature
;      LOCATION: (B) LOCATION 1...358
;      SEQUENCE DESCRIPTION: SEQ ID NO: 4677:
US-09-107-532A-4677
```

```
Query Match      90.0%;  Score 27;  DB 4;  Length 358;
Best Local Similarity  83.3%;  Pred. No. 4.4e+02;
Matches      5;  Conservative      1;  Mismatches      0;  Indels      0;  Gaps      0;
```

```
Qy      1 LLNNMR 6
      ||||:|
Db      104 LLNNLR 109
```

RESULT 10

```
US-09-248-796A-15218
; Sequence 15218, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA
ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
```

; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 15218  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-15218

Query Match 90.0%; Score 27; DB 4; Length 475;  
Best Local Similarity 83.3%; Pred. No. 5.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||:|  
Db 261 LLNNLR 266

RESULT 11

US-09-543-681A-6512  
; Sequence 6512, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS  
MIRABILIS FOR  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6512  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6512

Query Match 90.0%; Score 27; DB 4; Length 637;  
Best Local Similarity 83.3%; Pred. No. 7.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|:||||  
Db 293 LVNNMR 298

RESULT 12

US-09-489-039A-7465  
; Sequence 7465, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 7465  
; LENGTH: 726  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7465

Query Match 90.0%; Score 27; DB 4; Length 726;  
Best Local Similarity 83.3%; Pred. No. 8.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||:|  
Db 349 LLNNLR 354

RESULT 13

US-09-328-352-4373  
; Sequence 4373, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO. 4373  
; LENGTH: 814  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4373

Query Match 90.0%; Score 27; DB 4; Length 814;  
Best Local Similarity 83.3%; Pred. No. 9.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
:||||  
Db 335 VLNNMR 340

RESULT 14

US-09-749-340-6  
; Sequence 6, Application US/09749340  
; Patent No. 6593087  
; GENERAL INFORMATION:  
; APPLICANT: Prichard, Roger K.  
; APPLICANT: Xu, Ming  
; APPLICANT: Ribeiro, Ana Paula  
; APPLICANT: Blackhall, William J.  
; APPLICANT: Beech, Robin N.

; APPLICANT: Molento, Marcelo  
 ; APPLICANT: Liu, Hao Yuan  
 ; TITLE OF INVENTION: Methods for Detecting and Reversing  
 ; TITLE OF INVENTION: Resistance to Macrocyclic Lactone Compounds  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: American Home Products Corporation  
 ; STREET: One Campus Drive  
 ; CITY: Parsippany  
 ; STATE: New Jersey  
 ; COUNTRY: United States  
 ; ZIP: 07054  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/749,340  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/067,676  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Renda, Barbara L.  
 ; REGISTRATION NUMBER: 27,626  
 ; REFERENCE/DOCKET NUMBER: 33,333-01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 973-683-2153  
 ; TELEFAX: 973-683-4109  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1275 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-09-749-340-6

Query Match 90.0%; Score 27; DB 4; Length 1275;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 :|||||  
 Db 385 VLNNMR 390

RESULT 15  
 US-09-107-532A-4450  
 ; Sequence 4450, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO



; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND  
THERAPEUTICS

; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 4450:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 65 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES

; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium

; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...65  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4450:  
US-09-107-532A-4450

Query Match 86.7%; Score 26; DB 4; Length 65;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||:|  
Db 4 LLNNVR 9

Search completed: February 10, 2005, 16:02:10  
Job time : 14.4366 secs

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:49:10 ; Search time 35.9155 Seconds  
(without alignments)  
54.586 Million cell updates/sec

Title: US-10-067-484-9  
Perfect score: 30  
Sequence: 1 LLNNMR 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result                      %  
Query

No.	Score	Match	Length	DB	ID	Description
1	30	100.0	6	14	US-10-067-484-9	Sequence 9, Appli
2	30	100.0	6	14	US-10-067-620-9	Sequence 9, Appli
3	30	100.0	49	16	US-10-767-701-62883	Sequence 62883, A
4	30	100.0	459	16	US-10-473-574-1	Sequence 1, Appli
5	30	100.0	593	15	US-10-435-696-50	Sequence 50, Appl
6	30	100.0	618	9	US-09-925-300-1381	Sequence 1381, Ap
7	30	100.0	875	16	US-10-437-963-201180	Sequence 201180,
8	30	100.0	1058	15	US-10-424-599-224979	Sequence 224979,
9	30	100.0	1083	10	US-09-835-976B-14	Sequence 14, Appl
10	30	100.0	1106	16	US-10-322-696-39	Sequence 39, Appl
11	30	100.0	1243	15	US-10-425-114-63128	Sequence 63128, A
12	28	93.3	721	15	US-10-296-115-951	Sequence 951, App
13	28	93.3	1099	10	US-09-835-976B-10	Sequence 10, Appl
14	28	93.3	1116	10	US-09-835-976B-12	Sequence 12, Appl
15	28	93.3	1116	15	US-10-276-774-2161	Sequence 2161, Ap
16	28	93.3	1135	10	US-09-835-976B-4	Sequence 4, Appli
17	28	93.3	1135	10	US-09-835-976B-6	Sequence 6, Appli
18	28	93.3	1150	10	US-09-835-976B-8	Sequence 8, Appli
19	28	93.3	1150	10	US-09-835-976B-16	Sequence 16, Appl
20	27	90.0	40	15	US-10-424-599-277219	Sequence 277219,
21	27	90.0	121	16	US-10-437-963-106555	Sequence 106555,
22	27	90.0	172	15	US-10-282-122A-57513	Sequence 57513, A
23	27	90.0	199	16	US-10-437-963-132298	Sequence 132298,
24	27	90.0	202	15	US-10-424-599-213143	Sequence 213143,
25	27	90.0	306	14	US-10-156-761-9786	Sequence 9786, Ap
26	27	90.0	313	15	US-10-424-599-157386	Sequence 157386,
27	27	90.0	322	14	US-10-032-201B-221	Sequence 221, App
28	27	90.0	523	15	US-10-369-493-8418	Sequence 8418, Ap
29	27	90.0	612	15	US-10-282-122A-78158	Sequence 78158, A
30	27	90.0	618	15	US-10-282-122A-68453	Sequence 68453, A
31	27	90.0	798	15	US-10-282-122A-45056	Sequence 45056, A
32	27	90.0	847	14	US-10-177-293-498	Sequence 498, App
33	27	90.0	850	14	US-10-128-714-3534	Sequence 3534, Ap
34	27	90.0	883	14	US-10-128-714-8534	Sequence 8534, Ap
35	27	90.0	936	15	US-10-369-493-3625	Sequence 3625, Ap
36	27	90.0	1094	14	US-10-032-585-7554	Sequence 7554, Ap
37	27	90.0	1210	14	US-10-128-714-3078	Sequence 3078, Ap
38	27	90.0	1210	16	US-10-437-963-131764	Sequence 131764,
39	27	90.0	1246	14	US-10-128-714-8078	Sequence 8078, Ap
40	27	90.0	1275	9	US-09-749-340-6	Sequence 6, Appli
41	27	90.0	1275	14	US-10-456-815-6	Sequence 6, Appli
42	27	90.0	1507	15	US-10-282-122A-51807	Sequence 51807, A
43	27	90.0	1774	16	US-10-437-963-107694	Sequence 107694,
44	27	90.0	2538	16	US-10-609-113-49	Sequence 49, Appl
45	27	90.0	2538	16	US-10-753-901-20	Sequence 20, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-067-484-9

; Sequence 9, Application US/10067484

; Publication No. US20030170763A1

; GENERAL INFORMATION:

```

; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; TITLE OF INVENTION: RAGWEED ALLERGENS
; FILE REFERENCE: 416272000200
; CURRENT APPLICATION NUMBER: US/10/067,484
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Ragweed
US-10-067-484-9

```

```

Query Match      100.0%; Score 30; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 LLNNMR 6
        |||||
Db      1 LLNNMR 6

```

#### RESULT 2

```

US-10-067-620-9
; Sequence 9, Application US/10067620
; Publication No. US20030180225A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; APPLICANT: Teuber, Suzanne S.
; TITLE OF INVENTION: WALNUT AND RYEGRASS ALLERGENS
; FILE REFERENCE: 416272003400
; CURRENT APPLICATION NUMBER: US/10/067,620
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Ragweed
US-10-067-620-9

```

```

Query Match      100.0%; Score 30; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 LLNNMR 6
        |||||
Db      1 LLNNMR 6

```

RESULT 3

US-10-767-701-62883

; Sequence 62883, Application US/10767701  
 ; Publication No. US20040172684A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated  
 With  
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
 ; FILE REFERENCE: 38-21(53535)B  
 ; CURRENT APPLICATION NUMBER: US/10/767,701  
 ; CURRENT FILING DATE: 2004-01-29  
 ; NUMBER OF SEQ ID NOS: 63128  
 ; SEQ ID NO 62883  
 ; LENGTH: 49  
 ; TYPE: PRT  
 ; ORGANISM: Sorghum bicolor  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 18066725.pep  
 US-10-767-701-62883

Query Match 100.0%; Score 30; DB 16; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 8 LLNNMR 13

RESULT 4

US-10-473-574-1

; Sequence 1, Application US/10473574  
 ; Publication No. US20040116670A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;  
 ; APPLICANT: TANG, Y. Tom; YUE, Henry;  
 ; APPLICANT: KHAN, Farrah A.; ISON, Craig H.;  
 ; APPLICANT: BAUGHN, Mariah R.; WARREN, Bridget A.;  
 ; APPLICANT: DUGGAN, Brendan M.; THANGAVELU, Kavitha;  
 ; APPLICANT: HONCHELL, Cynthia D.; AZIMZAI, Yalda;  
 ; APPLICANT: ELLIOTT, Vicki S.; BURFORD, Neil;  
 ; APPLICANT: DING, Li; YUE, Huibin;  
 ; APPLICANT: BECHA, Shanya; EMERLING, Brooke M.;  
 ; APPLICANT: RICHARDSON, Thomas W.; LEE, Soo Yeun;  
 ; APPLICANT: BANDMAN, Olga; LAL, Preeti G.;  
 ; APPLICANT: LEE, Sally; GIETZEN, Kimberly J.;  
 ; APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;  
 ; APPLICANT: LEE, Ernestine A.; SWARNAKAR, Anita;  
 ; APPLICANT: RING, Huijun Z.; JONES, Karen Anne  
 ; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS  
 ; FILE REFERENCE: PF-0918 USN  
 ; CURRENT APPLICATION NUMBER: US/10/473,574  
 ; CURRENT FILING DATE: 2003-09-29

```

; PRIOR APPLICATION NUMBER: PCT/US02/09288
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US 60/294,451
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/291,870
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/290,518
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/288,609
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/283,769
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/281,323
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/280,508
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 1
;   LENGTH: 459
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 6582721CD1
US-10-473-574-1

```

```

Query Match          100.0%; Score 30; DB 15; Length 459;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches      6; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      1 LLNNMR 6
        |||||
Db      256 LLNNMR 261

```

# RESULT 5

US-10-435-696-50

```

; Sequence 50, Application US/10435696
; Publication No. US20040018525A1
; GENERAL INFORMATION:
; APPLICANT: Wirtz, Ralph
; APPLICANT: Munnes, Marc
; APPLICANT: Kallabis, Harald
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS,
PROGNOSIS,
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA
; FILE REFERENCE: LeA 36 108
; CURRENT APPLICATION NUMBER: US/10/435,696
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: EP03003112.4
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: EP02010291.9
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50

```

; LENGTH: 593  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-435-696-50

Query Match 100.0%; Score 30; DB 15; Length 593;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 317 LLNNMR 322

RESULT 6

US-09-925-300-1381  
; Sequence 1381, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1381  
; LENGTH: 618  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (507)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (524)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (562)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1381

Query Match 100.0%; Score 30; DB 9; Length 618;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 328 LLNNMR 333

RESULT 7

US-10-437-963-201180  
 ; Sequence 201180, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules  
 Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 201180  
 ; LENGTH: 875  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_9657C.1.pep.  
 US-10-437-963-201180

Query Match 100.0%; Score 30; DB 16; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 213 LLNNMR 218

# RESULT 8

US-10-424-599-224979  
 ; Sequence 224979, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
 With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 224979  
 ; LENGTH: 1058  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:



; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_45187C.1.pep  
US-10-424-599-224979

Query Match 100.0%; Score 30; DB 15; Length 1058;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 61 LLNNMR 66

RESULT 9

US-09-835-976B-14

; Sequence 14, Application US/09835976B

; Publication No. US20030027983A1

; GENERAL INFORMATION:

; APPLICANT: Mount, David B.

; APPLICANT: Delpire, Eric

; APPLICANT: Gamba, Gerardo

; APPLICANT: Alfred L. George, Jr.

; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER  
NUCLEIC ACIDS AND

; TITLE OF INVENTION: POLYPEPTIDES AND

; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME

; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2

; CURRENT APPLICATION NUMBER: US/09/835,976B

; CURRENT FILING DATE: 2001-04-16

; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 1083

; TYPE: PRT

; ORGANISM: mouse

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (71)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (155)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (467)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (586)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (639)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1023)

; OTHER INFORMATION: Xaa=Leu or Ile  
US-09-835-976B-14

Query Match 100.0%; Score 30; DB 10; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 247 LLNNMR 252

RESULT 10

US-10-322-696-39

; Sequence 39, Application US/10322696  
; Publication No. US20040166490A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Malandro, Marc  
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER  
; FILE REFERENCE: 529452001200  
; CURRENT APPLICATION NUMBER: US/10/322,696  
; CURRENT FILING DATE: 2003-10-17  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 1106  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-322-696-39

Query Match 100.0%; Score 30; DB 16; Length 1106;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 270 LLNNMR 275

RESULT 11

US-10-425-114-63128

; Sequence 63128, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 63128  
; LENGTH: 1243  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY044A04\_FLI.pep  
US-10-425-114-63128

Query Match 100.0%; Score 30; DB 15; Length 1243;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||  
Db 246 LLNNMR 251

RESULT 12

US-10-296-115-951  
; Sequence 951, Application US/10296115  
; Publication No. US20040053248A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784PCT  
; CURRENT APPLICATION NUMBER: US/10/296,115  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 1478  
; SEQ ID NO 951  
; LENGTH: 721  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-296-115-951

Query Match 93.3%; Score 28; DB 15; Length 721;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
:|||  
Db 77 MLNNMR 82

RESULT 13

US-09-835-976B-10  
; Sequence 10, Application US/09835976B  
; Publication No. US20030027983A1  
; GENERAL INFORMATION:  
; APPLICANT: Mount, David B.  
; APPLICANT: Delpire, Eric  
; APPLICANT: Gamba, Gerardo  
; APPLICANT: Alfred L. George, Jr.

```

; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER
NUCLEIC ACIDS AND
; TITLE OF INVENTION: POLYPEPTIDES AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1099
; TYPE: PRT
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (66)
; OTHER INFORMATION: Xaa=Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (537)
; OTHER INFORMATION: Xaa=Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (974)
; OTHER INFORMATION: n=c or a, Xaa=Leu or Ile
US-09-835-976B-10

```

```

Query Match          93.3%; Score 28; DB 10; Length 1099;
Best Local Similarity 83.3%; Pred. No. 1.8e+03;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 LLNNMR 6
        :|||||
Db      263 MLNNMR 268

```

#### RESULT 14

```

US-09-835-976B-12
; Sequence 12, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER
NUCLEIC ACIDS AND
; TITLE OF INVENTION: POLYPEPTIDES AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1116
; TYPE: PRT

```

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (120)  
; OTHER INFORMATION: Xaa=Leu or Ile  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (358)  
; OTHER INFORMATION: Xaa=Leu or Ile  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (916)  
; OTHER INFORMATION: Xaa=Leu or Ile  
US-09-835-976B-12

Query Match 93.3%; Score 28; DB 10; Length 1116;  
Best Local Similarity 83.3%; Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6  
:|||||  
Db 226 MLNNMR 231

RESULT 15

US-10-276-774-2161  
; Sequence 2161, Application US/10276774  
; Publication No. US20040053245A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang, Y, Tom et al  
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-030  
; CURRENT APPLICATION NUMBER: US/10/276,774  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 2700  
; SOFTWARE: Custom  
; SEQ ID NO 2161  
; LENGTH: 1116  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-276-774-2161

Query Match 93.3%; Score 28; DB 15; Length 1116;  
Best Local Similarity 83.3%; Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6  
:|||||  
Db 280 MLNNMR 285

Search completed: February 10, 2005, 16:41:34

Job time : 36.9155 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 9.29577 Seconds  
(without alignments)  
62.104 Million cell updates/sec

Title: US-10-067-484-9  
Perfect score: 30  
Sequence: 1 LLNNMR 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	30	100.0	526	1	KRBOVI	keratin, 54K type
2	30	100.0	561	2	A31994	keratin 10, type I
3	30	100.0	569	1	KRMSE1	keratin, 59K type
4	30	100.0	570	2	S07330	keratin, epidermal
5	30	100.0	593	1	KRHU0	keratin 10, type I
6	28	93.3	126	2	S77058	hypothetical prote
7	28	93.3	351	2	JC2451	Cwnt-4 protein pre
8	28	93.3	352	2	A49146	developmental regu
9	28	93.3	429	2	A25145	keratin, 47K type
10	28	93.3	433	2	H82139	trigger factor [si
11	28	93.3	434	2	AF0383	Trigger factor [im
12	28	93.3	1014	2	T17275	hypothetical prote

13	28	93.3	1116	2	T31432	K-Cl cotransport p
14	28	93.3	2359	2	T03094	A-kinase anchor pr
15	27	90.0	187	2	T06581	probable deoxyribo
16	27	90.0	322	2	T40393	thioredoxin-disulf
17	27	90.0	426	2	D42519	A20R protein - vac
18	27	90.0	426	2	T37408	probable 49.1K pro
19	27	90.0	426	2	C72166	A22R protein - var
20	27	90.0	426	2	T28563	hypothetical prote
21	27	90.0	426	2	C36850	A21R protein - var
22	27	90.0	447	2	F97828	conserved GTP-bind
23	27	90.0	447	2	H71673	hypothetical prote
24	27	90.0	489	2	T39038	hypothetical prote
25	27	90.0	536	2	T05500	calcium-dependent
26	27	90.0	571	2	T00835	calcium-dependent
27	27	90.0	612	2	B84936	RNA polymerase sig
28	27	90.0	612	2	AF0079	RNA polymerase sig
29	27	90.0	617	1	RNJV7A	transcription init
30	27	90.0	720	1	G64836	probable membrane
31	27	90.0	720	2	D90759	hypothetical prote
32	27	90.0	720	2	B85623	hypothetical prote
33	27	90.0	950	2	T44422	oxoglutarate dehyd
34	27	90.0	966	2	E84053	penicillin-binding
35	27	90.0	1275	2	T31073	multidrug resistan
36	27	90.0	1507	2	D97106	large chain of NAD
37	27	90.0	1979	2	C71622	hypothetical prote
38	27	90.0	2052	2	T37711	probable n-end-rec
39	27	90.0	4092	1	S38128	dynein heavy chain
40	26	86.7	174	2	G69824	conserved hypothet
41	26	86.7	224	2	D90235	hypothetical prote
42	26	86.7	242	2	T24034	hypothetical prote
43	26	86.7	264	2	B46314	hypothetical prote
44	26	86.7	265	2	G83240	probable permease
45	26	86.7	337	2	A53041	effector cell prot

#### ALIGNMENTS

##### RESULT 1

KRBOVI

keratin, 54K type I cytoskeletal - bovine

N;Alternate names: 54-kDa type I keratin; cytokeratin VIb

C;Species: Bos primigenius taurus (cattle)

C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004

C;Accession: A02941; S07262

R;Rieger, M.; Jorcano, J.L.; Franke, W.W.

EMBO J. 4, 2261-2267, 1985

A;Title: Complete sequence of a bovine type I cytokeratin gene: conserved and variable intron positions in genes of polypeptides of the same cytokeratin subfamily.

A;Reference number: A02941; MUID:86081734; PMID:2416562

A;Accession: A02941

A;Molecule type: DNA

A;Residues: 1-526 <RIE>

A;Cross-references: UNIPROT:P06394; GB:X02870; NID:g478; PIDN:CAA26626.1;

PID:g479

R;Jorcano, J.L.; Rieger, M.; Franz, J.K.; Schiller, D.L.; Moll, R.; Franke, W.W.

J. Mol. Biol. 179, 257-281, 1984

A;Title: Identification of two types of keratin polypeptides within the acidic cytokeratin subfamily I.

A;Reference number: S07262; MUID:85058191; PMID:6209405

A;Accession: S07262

A;Molecule type: mRNA

A;Residues: 281-466, 'PAAATAAEVQRWRPRRXFRRQYG', 490-491, 494, 'PVAVARRRK', 504-526 <JOR>

A;Cross-references: EMBL:X01460

A;Note: this sequence has been revised in reference A02941

C;Genetics:

A;Gene: cytokeratin VIb

A;Introns: 190/3; 218/2; 270/3; 324/3; 366/3; 439/2; 525/2

C;Superfamily: cytoskeletal keratin

C;Keywords: coiled coil; intermediate filament

F;1-127/Domain: head <HED>

F;1-127/Region: E1 and V1 subdomains

F;128-441/Domain: rod <ROD>

F;128-162/Region: coil 1A

F;163-176/Region: linker 1

F;177-277/Region: coil 1B

F;278-293/Region: linker 12

F;294-312/Region: coil 2A

F;313-320/Region: linker 2

F;321-441/Region: coil 2B

F;379/Region: stutter

F;442-526/Domain: tail <END>

F;442-526/Region: V2 and E2 subdomains

Query Match 100.0%; Score 30; DB 1; Length 526;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6

|||||

Db 298 LLNNMR 303

## RESULT 2

A31994

keratin 10, type I, epidermal - human

N;Alternate names: keratin 10 V2 variant 117

C;Species: Homo sapiens (man)

C;Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 09-Jul-2004

C;Accession: A31994; A38182

R;Zhou, X.M.; Idler, W.W.; Steven, A.C.; Roop, D.R.; Steinert, P.M.

J. Biol. Chem. 263, 15584-15589, 1988

A;Title: The complete sequence of the human intermediate filament chain keratin 10. Subdomainal divisions and model for folding of end domain sequences.

A;Reference number: A31994; MUID:89008465; PMID:2459124

A;Accession: A31994

A;Molecule type: mRNA

A;Residues: 1-561 <ZHO>

A;Cross-references: UNIPROT:Q14664; GB:J04029; NID:g623408; PIDN:AAA60544.1; PID:g623409

R;Korge, B.P.; Gan, S.Q.; McBride, O.W.; Mischke, D.; Steinert, P.M.

Proc. Natl. Acad. Sci. U.S.A. 89, 910-914, 1992



A;Title: Extensive size polymorphism of the human keratin 10 chain resides in the C-terminal V2 subdomain due to variable numbers and sizes of glycine loops.  
A;Reference number: A38182; MUID:92141228; PMID:1371013  
A;Accession: A38182  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 445-561 <KOR>  
A;Note: sequence extracted from NCBI backbone (NCBIP:79433)  
C;Superfamily: cytoskeletal keratin  
C;Keywords: coiled coil

Query Match 100.0%; Score 30; DB 2; Length 561;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 310 LLNNMR 315

### RESULT 3

#### KRMSE1

keratin, 59K type I cytoskeletal - mouse

N;Alternate names: 59-kDa type I keratin

C;Species: Mus musculus (house mouse)

C;Date: 15-Nov-1984 #sequence\_revision 04-Dec-1986 #text\_change 09-Jul-2004

C;Accession: A02940

R;Krieg, T.M.; Schafer, M.P.; Cheng, C.K.; Filpula, D.; Flaherty, P.; Steinert, P.M.; Roop, D.R.

J. Biol. Chem. 260, 5867-5870, 1985

A;Title: Organization of a type I keratin gene. Evidence for evolution of intermediate filaments from a common ancestral gene.

A;Reference number: A02940; MUID:85207552; PMID:2581944

A;Accession: A02940

A;Molecule type: DNA

A;Residues: 1-569 <KRI>

A;Cross-references: UNIPROT:P02535; GB:L00193; GB:K00391; NID:g198625;

PIDN:AAA39391.1; PID:g387397

A;Note: initiator Met not shown

A;Note: the authors translated the codon GAG for residue 41 as Gly

C;Comment: Fourier analysis has identified a 7-residue repeating pattern (heptad) between residues 138-454, in which the first and fourth residues are apolar and the remainder are polar. This region forms a stable alpha-helical coiled coil but is interrupted by three short regions with beta conformation.

C;Comment: Most of the introns of the gene encoding this protein are located within the regions predicted to form alpha helices; although they do not demarcate structural domains, they do interrupt the sequence at or near the beginning of heptad repeats. Several of these sites are conserved in genes encoding type II and III intermediate filaments.

C;Comment: The amino and carboxyl ends are rich in glycine, serine, and aromatic residues arranged in short, tandem repeats characteristic of cytoskeletal keratins.

C;Genetics:

A;Introns: 206/3; 234/2; 286/3; 340/3; 382/3; 455/2; 568/2

C;Superfamily: cytoskeletal keratin

C;Keywords: coiled coil; intermediate filament

F;1-143/Domain: head <HED>

F;1-143/Region: E1 and V1 subdomains  
F;144-457/Domain: rod <ROD>  
F;144-178/Region: coil 1A  
F;179-192/Region: linker 1  
F;193-293/Region: coil 1B  
F;294-309/Region: linker 12  
F;310-328/Region: coil 2A  
F;329-336/Region: linker 2  
F;337-457/Region: coil 2B  
F;395/Region: stutter  
F;458-569/Domain: tail <END>  
F;458-569/Region: V2 and E2 subdomains

Query Match 100.0%; Score 30; DB 1; Length 569;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 314 LLNNMR 319

RESULT 4

S07330

keratin, epidermal - mouse

C;Species: Mus musculus (house mouse)

C;Date: 01-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C;Accession: S07330

R;Steinert, P.M.; Rice, R.H.; Roop, D.R.; Trus, B.L.; Steven, A.C.

Nature 302, 794-800, 1983

A;Title: Complete amino acid sequence of a mouse epidermal keratin subunit and implications for the structure of intermediate filaments.

A;Reference number: S07330; MUID:83192464; PMID:6188955

A;Accession: S07330

A;Molecule type: mRNA

A;Residues: 1-570 <STE>

A;Cross-references: UNIPROT:P02535; EMBL:V00830; NID:g52786; PIDN:CAA24214.1;

PID:g52787

A;Note: the authors translated the codon TTC for residue 25 as Ser, TTC for residue 29 as Ser, and GGT for residue 523 as Ser

C;Superfamily: cytoskeletal keratin

C;Keywords: coiled coil; intermediate filament

Query Match 100.0%; Score 30; DB 2; Length 570;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 314 LLNNMR 319

RESULT 5

KRHU0

keratin 10, type I, cytoskeletal - human

N;Alternate names: cytokeratin 10

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
 C;Accession: S02158; C38182; B38182; PC1102; S14666; S14669  
 R;Rieger, M.; Franke, W.W.  
 J. Mol. Biol. 204, 841-856, 1988  
 A;Title: Identification of an orthologous mammalian cytokeratin gene. High degree of intron sequence conservation during evolution of human cytokeratin 10.  
 A;Reference number: S02158; MUID:89125611; PMID:2464696  
 A;Accession: S02158  
 A;Molecule type: DNA  
 A;Residues: 1-593 <RIE>  
 A;Cross-references: UNIPROT:P13645; EMBL:X14487; NID:g28316; PIDN:CAA32649.1; PID:g28317  
 A;Experimental source: clone lambda-KH10-5  
 R;Korge, B.P.; Gan, S.Q.; McBride, O.W.; Mischke, D.; Steinert, P.M.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 910-914, 1992  
 A;Title: Extensive size polymorphism of the human keratin 10 chain resides in the C-terminal V2 subdomain due to variable numbers and sizes of glycine loops.  
 A;Reference number: A38182; MUID:92141228; PMID:1371013  
 A;Accession: C38182  
 A;Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 452-593 <KOR1>  
 A;Cross-references: PIDN:AAB21315.1; PID:g244509  
 A;Note: sequence extracted from NCBI backbone (NCBIP:79427)  
 A;Accession: B38182  
 A;Status: preliminary; not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 452-463, 'P', 465-507, 'Y', 523-593 <KOR2>  
 A;Cross-references: PIDN:AAB21314.1; PID:g244508  
 A;Note: sequence extracted from NCBI backbone (NCBIP:79431)  
 R;Tkachenko, A.V.; Buchman, V.L.; Bliskovsky, V.V.; Shvets, Y.P.; Kisselev, L.L.  
 Gene 116, 245-251, 1992  
 A;Title: Exons I and VII of the gene (Ker10) encoding human keratin 10 undergo structural rearrangements within repeats.  
 A;Reference number: PC1102; MUID:92339897; PMID:1378806  
 A;Accession: PC1102  
 A;Molecule type: mRNA  
 A;Residues: 'G', 198-407, 'Q', 409-450, 'G', 452-486, 491-524, 534-593 <TKA>  
 A;Cross-references: GB:M77663; NID:g186628; PIDN:AAA59199.1; PID:g186629  
 A;Experimental source: embryonic skin, clone HK51  
 R;Darmon, M.Y.; Semat, A.; Darmon, M.C.; Vasseur, M.  
 Mol. Biol. Rep. 12, 277-283, 1987  
 A;Title: Sequence of a cDNA encoding human keratin No 10 selected according to structural homologies of keratins and their tissue-specific expression.  
 A;Reference number: S14666; MUID:88122104; PMID:2448602  
 A;Accession: S14666  
 A;Molecule type: mRNA  
 A;Residues: 130-278, 'YV', 281-311, 'I', 313-339, 'V', 341-373, 'R', 375-407, 'Q', 409-459, 'RS', 462-476, 'T', 478-481, 'T', 483-502, 'T', 504-507, 'T', 509-512, 'LRGELHGGHAH', 535, 'T', 537-542, 'N', 544-550, 'LR', 553, 'RH', 556-579, 'P', 581-593 <DAR1>  
 A;Cross-references: EMBL:M19156; NID:g186769  
 A;Note: the sequence from Fig. 3 is inconsistent with the nucleotide sequence from Fig. 2 in having additional residues Ile-Lys-Ile-Arg-Leu after 442-Leu  
 R;Darmon, M.Y.; Semat, A.; Darmon, M.C.; Vasseur, M.  
 submitted to the EMBL Data Library, May 1988  
 A;Reference number: S14667

A;Accession: S14669  
 A;Molecule type: mRNA  
 A;Residues: 130-278, 'YV', 281-311, 'I', 313-339, 'V', 341-373, 'R', 375-407, 'Q', 409-459, 'RS', 462-476, 'T', 478-481, 'T', 483-502, 'T', 504-507, 'T', 509-512, 'LRGELHGGHAH', 535, 'T', 537-542, 'N', 544-550, 'LR', 553, 'RH', 556-593 <DAR2>  
 A;Cross-references: EMBL:M19156; NID:g186769; PIDN:AAA59468.1; PID:g307086  
 A;Note: the translated sequence in GenBank entry HUMKRT10A, release 111.0, differs from the published sequence in beginning incorrectly at the internal 21-Met AUG codon  
 C;Genetics:  
 A;Gene: GDB:KRT10; KPP  
 A;Cross-references: GDB:118828; OMIM:148080  
 A;Map position: 17q12-17q21  
 A;Introns: 209/3; 237/2; 289/3; 343/3; 385/3; 458/2; 592/3  
 A;Note: this gene encodes variants with considerable length polymorphism  
 A;Note: mutations in this gene can cause epidermolytic hyperkeratosis and keratosis palmaris et plantaris  
 C;Complex: heterotetramer of two type I and two type II proteins, usually keratin 1 (see PIR:KRHU2)  
 C;Superfamily: cytoskeletal keratin  
 C;Keywords: coiled coil; heterotetramer; intermediate filament; polymorphism  
 F;1-145/Domain: head <HEA>  
 F;146-456/Domain: helical rod #status predicted <ROD>  
 F;457-593/Domain: tail <TAI>

Query Match 100.0%; Score 30; DB 1; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 317 LLNNMR 322

# RESULT 6

S77058

hypothetical protein sll0668 - *Synechocystis* sp. (strain PCC 6803)

C;Species: *Synechocystis* sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

C;Accession: S77058

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.;  
 Miyajima, N.; Hirose, M.; Sugiura, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.;  
 Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.;  
 Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium  
*Synechocystis* sp. PCC6803. II. Sequence determination of the entire genome and  
 assignment of potential protein-coding regions.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S77058

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-126 <KAN>

A;Cross-references: UNIPROT:Q55973; EMBL:D64005; GB:AB001339; NID:g1001779;  
 PIDN:BAA10750.1; PID:d1011401; PID:g1006597

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 93.3%; Score 28; DB 2; Length 126;  
Best Local Similarity 83.3%; Pred. No. 24;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
:|||||  
Db 64 ILNNMR 69

RESULT 7

JC2451

Cwnt-4 protein precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 21-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 03-May-1996

C;Accession: JC2451

R;Yoshioka, H.; Ohuchi, H.; Nohno, T.; Fujiwara, A.; Tanda, N.; Kawakami, Y.; Noji, S.

Biochem. Biophys. Res. Commun. 203, 1581-1588, 1994

A;Title: Regional expression of the Cwnt-4 gene in developing chick central nervous system in relationship to the diencephalic neuromere D2 and a dorsal domain of the spinal cord.

A;Reference number: JC2451; MUID:95032034; PMID:7945308

A;Accession: JC2451

A;Molecule type: mRNA

A;Residues: 1-351 <YOS>

C;Comment: This protein is involved in segmentation of forebrain into the neuromere D2 and in differentiation of the dorsal region of the spinal cord.

C;Genetics:

A;Gene: Cwnt-4

C;Superfamily: int-1 transforming protein

C;Keywords: glycoprotein

F;1-36/Domain: signal sequence #status predicted <SIG>

F;37-351/Product: Cwnt-4 protein #status predicted <MAT>

F;88,297/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.3%; Score 28; DB 2; Length 351;  
Best Local Similarity 83.3%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
:|||||  
Db 198 ILNNMR 203

RESULT 8

A49146

developmental regulator wnt-4 - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Nov-1995

C;Accession: A49146; C49764

R;McGrew, L.L.; Otte, A.P.; Moon, R.T.

Development 115, 463-473, 1992

A;Title: Analysis of Xwnt-4 in embryos of Xenopus laevis: a Wnt family member expressed in the brain and floor plate.

A;Reference number: A49146; MUID:93048810; PMID:1425335  
 A;Accession: A49146  
 A;Status: preliminary; not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-352 <MCG>  
 A;Experimental source: neurula stage embryos  
 A;Note: sequence extracted from NCBI backbone (NCBIP:117112)  
 R;Christian, J.L.; Gavin, B.J.; McMahon, A.P.; Moon, R.T.  
 Dev. Biol. 143, 230-234, 1991  
 A;Title: Isolation of cDNAs partially encoding four Xenopus Wnt-1/int-1-related proteins and characterization of their transient expression during embryonic development.  
 A;Reference number: A49764; MUID:91122437; PMID:1991549  
 A;Accession: C49764  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 262-295, 'H', 297-335 <CHR>  
 A;Cross-references: GB:M55055  
 C;Superfamily: int-1 transforming protein

Query Match 93.3%; Score 28; DB 2; Length 352;  
 Best Local Similarity 83.3%; Pred. No. 76;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 :|||||  
 Db 199 ILNNMR 204

RESULT 9  
 A25145  
 keratin, 47K type I cytoskeletal, larval - African clawed frog  
 N;Alternate names: XK81  
 C;Species: Xenopus laevis (African clawed frog)  
 C;Date: 22-Jul-1987 #sequence\_revision 22-Jul-1987 #text\_change 09-Jul-2004  
 C;Accession: A25145  
 R;Jonas, E.; Sargent, T.D.; Dawid, I.B.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 5413-5417, 1985  
 A;Reference number: A25145; MUID:85270534; PMID:2410923  
 A;Accession: A25145  
 A;Molecule type: mRNA  
 A;Residues: 1-429 <JON>  
 A;Cross-references: UNIPROT:P08777; GB:M11940; NID:g214565; PIDN:AAA49894.1;  
 PID:g214566  
 C;Superfamily: cytoskeletal keratin  
 C;Keywords: coiled coil

Query Match 93.3%; Score 28; DB 2; Length 429;  
 Best Local Similarity 83.3%; Pred. No. 95;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 :|||||  
 Db 242 ILNNMR 247

RESULT 10

H82139

trigger factor [similarity] - *Vibrio cholerae* (strain N16961 serogroup O1)

N;Contains: peptidylprolyl isomerase (EC 5.2.1.8), ribosome-bound

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C;Accession: H82139

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.;  
Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill,  
S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.;  
Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.;  
Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.;  
Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: H82139

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-433 <HEI>

A;Cross-references: UNIPROT:Q9KQS5; GB:AE004268; GB:AE003852; NID:g9656456;

PIDN:AAF95071.1; GSPDB:GN00126; TIGR:VC1923

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1923

A;Map position: 1

C;Superfamily: FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor);

BKEP-type peptidylprolyl isomerase homology

C;Keywords: cis-trans-isomerase

Query Match 93.3%; Score 28; DB 2; Length 433;

Best Local Similarity 83.3%; Pred. No. 96;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6

|:||||

Db 394 LMNNMR 399

RESULT 11

AF0383

Trigger factor [imported] - *Yersinia pestis* (strain CO92)

C;Species: *Yersinia pestis*

C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AF0383

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.;  
Prentice, M.B.; Sebaihia, M.; James, K.D.; Churcher, C.; Mungall, K.L.; Baker,  
S.; Basham, D.; Bentley, S.D.; Brooks, K.; Cerdano-Tarraga, A.M.; Chillingworth,  
T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Feltwell, T.; Hamlin, N.;  
Holroyd, S.; Jagels, K.; Leather, S.; Karlyshev, A.V.; Moule, S.; Oyston,  
P.C.F.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
Whitehead, S.; Barrell, B.G.

Nature 413, 523-527, 2001

A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AF0383

A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-434 <KUR>  
A;Cross-references: UNIPROT:Q8ZC64; GB:AL590842; PIDN:CAC92393.1; PID:g15981096;  
GSPDB:GN00175  
C;Genetics:  
A;Gene: tig  
C;Superfamily: FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor);  
BKBP-type peptidylprolyl isomerase homology

Query Match 93.3%; Score 28; DB 2; Length 434;  
Best Local Similarity 83.3%; Pred. No. 96;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|:||||  
Db 394 LMNNMR 399

RESULT 12

T17275  
hypothetical protein DKFZp434D2135.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T17275  
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A;Reference number: Z18723  
A;Accession: T17275  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1014 <POU>  
A;Cross-references: UNIPROT:Q9UHW9; EMBL:AL117500  
A;Experimental source: adult testis; clone DKFZp434D2135  
C;Genetics:  
A;Note: DKFZp434D2135.1

Query Match 93.3%; Score 28; DB 2; Length 1014;  
Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
:|||||  
Db 178 MLNNMR 183

RESULT 13

T31432  
K-Cl cotransport protein 2, furosemide-sensitive - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T31432  
R;Payne, J.A.; Stevenson, T.J.; Donaldson, L.F.  
J. Biol. Chem. 271, 16245-16252, 1996  
A;Title: Molecular characterization of a putative K-Cl cotransporter in rat  
brain: a neuronal-specific isoform.  
A;Reference number: Z21031; MUID:96279171; PMID:8663311  
A;Accession: T31432



A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-1116 <PAY>  
A;Cross-references: UNIPROT:Q63633; EMBL:U55816; NID:g1403708; PID:g1403709;  
PIDN:AAC52635.1  
A;Experimental source: strain Sprague Dawley; clone ERB10; 5ERB12; brain  
C;Genetics:  
A;Gene: KCC2  
C;Keywords: transmembrane protein

Query Match 93.3%; Score 28; DB 2; Length 1116;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
:|||||  
Db 226 MLNNMR 231

RESULT 14

T03094

A-kinase anchor protein DAKAP550 - fruit fly (*Drosophila melanogaster*)  
(fragment)

C;Species: *Drosophila melanogaster*

C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C;Accession: T03094

R;Han, J.D.; Baker, N.E.; Rubin, C.S.

J. Biol. Chem. 272, 26611-26619, 1997

A;Title: Molecular characterization of a novel A kinase anchor protein from *drosophila melanogaster*.

A;Reference number: Z14835; MUID:97476266; PMID:9334242

A;Accession: T03094

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-2359 <HAN>

A;Cross-references: UNIPROT:Q9W4E2; EMBL:AF003622; NID:g2393879;

PIDN:AAB83959.1; PID:g2393880

A;Experimental source: strain Canton S

C;Genetics:

A;Cross-references: FlyBase:FBgn0021748

A;Map position: X

Query Match 93.3%; Score 28; DB 2; Length 2359;  
Best Local Similarity 83.3%; Pred. No. 6.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|:|||||  
Db 2179 LMNNMR 2184

RESULT 15

T06581

probable deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - tomato (fragment)

N;Alternate names: cryptochrome 2; DNA photolyase; photoreactivating enzyme

C;Species: *Lycopersicon esculentum* (tomato)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T06581  
R;Kolukisaoglu, U.; Schneider-Poetsch, H.A.  
submitted to the EMBL Data Library, August 1997  
A;Description: The evolution of cryptochrome genes.  
A;Reference number: Z15773  
A;Accession: T06581  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-187 <KOL>  
A;Cross-references: UNIPROT:O24020; EMBL:AJ000695; PIDN:CAA04247.1  
C;Genetics:  
A;Gene: CYR2  
A;Note: Intron positions not resolved (incomplete sequence)  
C;Superfamily: deoxyribodipyrimidine photo-lyase  
C;Keywords: carbon-carbon lyase

Query Match 90.0%; Score 27; DB 2; Length 187;  
Best Local Similarity 83.3%; Pred. No. 64;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||:|  
Db 172 LLNNLR 177

Search completed: February 10, 2005, 15:59:36  
Job time : 10.2958 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 43.7746 Seconds  
(without alignments)  
70.188 Million cell updates/sec

Title: US-10-067-484-9  
Perfect score: 30  
Sequence: 1 LLNNMR 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	30	100.0		334	2	Q9JKB4	Q9jkb4 mus musculu
2	30	100.0		383	2	Q6IFW7	Q6ifw7 rattus norv
3	30	100.0		423	2	Q6DDD5	Q6ddd5 xenopus lae
4	30	100.0		434	2	Q6D828	Q6d828 erwinia car
5	30	100.0		440	2	Q6BH13	Q6bh13 debaryomyce
6	30	100.0		446	2	Q8VCW2	Q8vcw2 mus musculu
7	30	100.0		446	2	Q6IFX0	Q6ifx0 rattus norv
8	30	100.0		448	2	Q9Z320	Q9z320 mus musculu
9	30	100.0		449	2	Q6IFW8	Q6ifw8 rattus norv
10	30	100.0		450	2	Q7Z3Z0	Q7z3z0 homo sapien
11	30	100.0		450	2	Q6R650	Q6r650 capra hircu
12	30	100.0		450	2	Q9BGM5	Q9bgm5 ovis aries
13	30	100.0		459	2	Q7Z3Y8	Q7z3y8 homo sapien
14	30	100.0		460	2	Q6R649	Q6r649 capra hircu
15	30	100.0		461	2	Q6C2Y6	Q6c2y6 yarrowia li
16	30	100.0		462	2	Q8BUX3	Q8bux3 mus musculu
17	30	100.0		462	2	Q9D637	Q9d637 mus musculu
18	30	100.0		464	2	Q7Z3Y7	Q7z3y7 homo sapien
19	30	100.0		468	2	Q7Z3Y9	Q7z3y9 homo sapien
20	30	100.0		486	2	Q6ZP84	Q6zp84 homo sapien
21	30	100.0		510	2	Q8BV09	Q8bv09 mus musculu
22	30	100.0		526	1	K1CJ_BOVIN	P06394 bos taurus
23	30	100.0		526	2	Q6IFW6	Q6ifw6 rattus norv
24	30	100.0		528	2	Q8C6Y2	Q8c6y2 mus musculu
25	30	100.0		561	2	Q14664	Q14664 homo sapien
26	30	100.0		561	2	Q8BVU3	Q8bvu3 mus musculu
27	30	100.0		561	2	Q9CXH6	Q9cxh6 mus musculu
28	30	100.0		568	2	Q6EIZ0	Q6eiz0 canis famil
29	30	100.0		569	1	K1CJ_MOUSE	P02535 mus musculu
30	30	100.0		584	2	Q8N175	Q8n175 homo sapien
31	30	100.0		593	1	K1CJ_HUMAN	P13645 homo sapien
32	30	100.0		716	2	Q8VXB0	Q8vxb0 oryza sativ
33	30	100.0		778	2	Q652J4	Q652j4 oryza sativ
34	30	100.0		872	2	Q6G2C7	Q6g2c7 bartonella
35	30	100.0		1083	1	S127_MOUSE	Q9wvl3 mus musculu
36	30	100.0		1091	2	Q6KAS8	Q6kas8 mus musculu
37	30	100.0		1528	2	Q6FSK0	Q6fsk0 candida gla
38	30	100.0		3277	2	Q7Q1I0	Q7q1i0 anopheles g
39	28	93.3		126	2	Q55973	Q55973 synechocyst
40	28	93.3		171	2	Q6M0W8	Q6m0w8 methanococc
41	28	93.3		171	2	Q6IU10	Q6iu10 raja erinac
42	28	93.3		351	1	WNT4_CHICK	P49337 gallus gall
43	28	93.3		351	1	WNT4_XENLA	P49338 xenopus lae
44	28	93.3		375	2	Q7ZYN9	Q7zyn9 xenopus lae
45	28	93.3		425	2	Q80B07	Q80b07 orf virus.

## ALIGNMENTS

## RESULT 1

Q9JKB4

ID Q9JKB4 PRELIMINARY; PRT; 334 AA.  
 AC Q9JKB4;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Epidermal keratin 10 (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ; TISSUE=Liver;  
 RA Reichelt J., Magin T.M.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF245658; AAF65456.1; -.  
 DR HSSP; P08670; 1GK7.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 KW Keratin.  
 FT NON\_TER 334 334  
 SQ SEQUENCE 334 AA; 34983 MW; AD59C2F193A081E1 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 308 LLNNMR 313

## RESULT 2

Q6IFW7

ID Q6IFW7 PRELIMINARY; PRT; 383 AA.  
 AC Q6IFW7;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Type I keratin KA41.  
 GN Name=Ka41;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=BN;  
 RX MEDLINE=23257644; PubMed=15085952;  
 RA Hesse M., Zimek A., Weber K., Magin T.M.;  
 RT "Comprehensive analysis of keratin gene clusters in humans and  
 RT rodents.";  
 RL Eur. J. Cell Biol. 83:19-26(2004).  
 CC -!- MISCELLANEOUS: The sequence shown here is derived from an  
 CC EMBL/GenBank/DDBJ third party annotation (TPA) entry.  
 DR EMBL; BK004031; DAA04465.1; -.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 KW Keratin.  
 SQ SEQUENCE 383 AA; 41325 MW; 9C43C09D0BC08EEA CRC64;

Query Match 100.0%; Score 30; DB 2; Length 383;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 256 LLNNMR 261

# RESULT 3

## Q6DDD5

ID Q6DDD5 PRELIMINARY; PRT; 423 AA.  
 AC Q6DDD5;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Krt16-prov protein.  
 GN Name=krt16-prov;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.  
 DR EMBL; BC077649; AAH77649.1; -.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 DR PROSITE; PS00226; IF; UNKNOWN\_1.  
 KW Intermediate filament.  
 SQ SEQUENCE 423 AA; 47915 MW; 3C86F6B4A6F1071F CRC64;

Query Match 100.0%; Score 30; DB 2; Length 423;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6  
 |||||  
 Db 234 LLNNMR 239

#### RESULT 4

Q6D828

ID Q6D828 PRELIMINARY; PRT; 434 AA.  
 AC Q6D828;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Trigger factor.  
 GN Name=tig; OrderedLocusNames=ECA1147;  
 OS *Erwinia carotovora* (subsp. *atroseptica*) (*Pectobacterium atrosepticum*).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; *Pectobacterium*.  
 CX NCBI\_TaxID=29471;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI 1043 / ATCC BAA-672;  
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;  
RA Bell K.S., Sebaihia M., Pritchard L., Holden M.T.G., Hyman L.J.,  
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,  
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,  
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,  
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,  
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;  
RT "Genome sequence of the enterobacterial phytopathogen *Erwinia*  
RT *carotovora* subsp. *atroseptica* and characterization of virulence  
RT factors.";   
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).  
DR EMBL; BX950851; CAG74057.1; -.  
DR GO; GO:0006457; P:protein folding; IEA.  
DR GO; GO:0015031; P:protein transport; IEA.  
DR InterPro; IPR001179; FKBP\_PPIase.  
DR InterPro; IPR008880; Trigger\_C.  
DR InterPro; IPR008881; Trigger\_N.  
DR InterPro; IPR005215; Trig\_fac.  
DR Pfam; PF00254; FKBP\_C; 1.  
DR Pfam; PF05698; Trigger\_C; 1.  
DR Pfam; PF05697; Trigger\_N; 1.  
DR PIRSF; PIRSF003095; Trigger\_factor; 1.  
DR TIGRFAMs; TIGR00115; tig; 1.  
DR PROSITE; PS50059; FKBP\_PPIASE; 1.  
KW Complete proteome.  
SQ SEQUENCE 434 AA; 48141 MW; 1EC3DBAFD72D3C04 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 394 LLNNMR 399

# RESULT 5

Q6BH13

ID Q6BH13 PRELIMINARY; PRT; 440 AA.  
AC Q6BH13;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Similar to CA6069|IPF4847 *Candida albicans* IPF4847.  
GN ORFNames=DEHA0G23474g;  
OS *Debaryomyces hansenii* CBS767.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.  
OX NCBI\_TaxID=284592;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBS767;  
RG Genolevures;  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,

RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swennene D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44 (2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; CR382139; CAG91018.1; -.  
 DR GO; GO:0030693; F:caspase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR006031; XYPPX.  
 DR Pfam; PF02162; XYPPX; 4.  
 SQ SEQUENCE 440 AA; 47835 MW; 0B0835A4AC0F7382 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 408 LLNNMR 413

# RESULT 6

## Q8VCW2

ID Q8VCW2 PRELIMINARY; PRT; 446 AA.  
 AC Q8VCW2;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE RIKEN cDNA 4631426H08.  
 GN Name=4631426H08Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Salivary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,



RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 FN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Salivary gland;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC018391; AAH18391.1; -.  
 DR HSSP; P08670; 1GK7.  
 DR MGD; MGI:1918060; 4631426H08Rik.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 SQ SEQUENCE 446 AA; 48921 MW; 6B502F612AC4BE28 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 446;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 247 LLNNMR 252

# RESULT 7

## Q6IFX0

ID Q6IFX0 PRELIMINARY; PRT; 446 AA.  
 AC Q6IFX0;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Type I keratin KA38.  
 GN Name=Ka38;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BN;  
 RX MEDLINE=23257644; PubMed=15085952;  
 RA Hesse M., Zimek A., Weber K., Magin T.M.;  
 RT "Comprehensive analysis of keratin gene clusters in humans and  
 RT rodents.";

RL Eur. J. Cell Biol. 83:19-26(2004).  
 CC -!- MISCELLANEOUS: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ third party annotation (TPA) entry..  
 DR EMBL; BK004028; DAA04462.1; -.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 KW Keratin.  
 SQ SEQUENCE 446 AA; 48943 MW; 3E78915CFA4C7CBD CRC64;

Query Match 100.0%; Score 30; DB 2; Length 446;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 247 LLNNMR 252

#### RESULT 8

Q9Z320

ID Q9Z320 PRELIMINARY; PRT; 448 AA.  
 AC Q9Z320;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE C29.  
 GN Name=Krt1-c29; Synonyms=c29;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL10J; TISSUE=Skin;  
 RX MEDLINE=99189238; PubMed=10087197; DOI=10.1006/geno.1998.5721;  
 RA Sato H., Koide T., Sagai T., Ishiguro S., Tamai M., Saitou N.,  
 RA Shiroishi T.;  
 RT "The genomic organization of type I keratin genes in mice.";  
 RL Genomics 56:303-309(1999).  
 DR EMBL; AB013607; BAA34228.1; -.  
 DR HSSP; P08670; 1GK7.  
 DR MGD; MGI:1339999; Krt1-c29.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 SQ SEQUENCE 448 AA; 49104 MW; 14CAF4EAC0828AEA CRC64;

Query Match 100.0%; Score 30; DB 2; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 246 LLNNMR 251

RESULT 9

Q6IFW8

ID Q6IFW8 PRELIMINARY; PRT; 449 AA.  
AC Q6IFW8;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Type I keratin KA40.  
GN Name=Ka40;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BN;  
RX MEDLINE=23257644; PubMed=15085952;  
RA Hesse M., Zimek A., Weber K., Magin T.M.;  
RT "Comprehensive analysis of keratin gene clusters in humans and  
rodents."  
RL Eur. J. Cell Biol. 83:19-26(2004).  
CC -!- MISCELLANEOUS: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.  
DR EMBL; BK004030; DAA04464.1; -.  
DR GO; GO:0005882; C:intermediate filament; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR001664; IF.  
DR InterPro; IPR002957; Keratin\_I.  
DR Pfam; PF00038; Filament; 1.  
DR PRINTS; PR01248; TYPE1KERATIN.  
KW Keratin.  
SQ SEQUENCE 449 AA; 49108 MW; FD863F380DDBAD1B CRC64;

Query Match 100.0%; Score 30; DB 2; Length 449;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 246 LLNNMR 251

RESULT 10

Q7Z3Z0

ID Q7Z3Z0 PRELIMINARY; PRT; 450 AA.  
AC Q7Z3Z0;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Type I inner root sheath specific keratin 25 irs1.  
GN Name=KRT25A;  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Human scalp;  
 RA Langbein L., Rogers M.A., Winter H., Praetzel S., Schweizer J.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AJ564204; CAD91904.1; -.  
 DR HSSP; P08670; 1GK7.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR Pfam; PF00038; Filament; 1.  
 KW Keratin.  
 SQ SEQUENCE 450 AA; 49318 MW; 12AB5BF74BA426FF CRC64;

Query Match 100.0%; Score 30; DB 2; Length 450;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLNNMR 6  
 |||||  
 Db 251 LLNNMR 256

# RESULT 11

Q6R650

ID Q6R650 PRELIMINARY; PRT; 450 AA.  
 AC Q6R650;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Type I keratin intermediate filament IRSa1.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yin J., Li J.Q., Zhou H.M.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY510111; AAS00518.1; -.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 KW Keratin.  
 SQ SEQUENCE 450 AA; 49423 MW; 084877188F3FFA90 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 450;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 251 LLNNMR 256

RESULT 12

Q9BGM5

ID Q9BGM5 PRELIMINARY; PRT; 450 AA.  
AC Q9BGM5;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Type I keratin intermediate filament IRSa1.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Corriedale;  
RA Bawden C.S., McLaughlan C.J., Nesci A., Rogers G.E.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF227759; AAK00222.1; -.  
DR HSSP; P08670; 1GK7.  
DR GO; GO:0005882; C:intermediate filament; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR001664; IF.  
DR InterPro; IPR002957; Keratin\_I.  
DR Pfam; PF00038; Filament; 1.  
DR PRINTS; PR01248; TYPE1KERATIN.  
KW Keratin.  
SQ SEQUENCE 450 AA; 49313 MW; 8F924BA92D3BB1D6 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 450;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
|||||  
Db 251 LLNNMR 256

RESULT 13

Q7Z3Y8

ID Q7Z3Y8 PRELIMINARY; PRT; 459 AA.  
AC Q7Z3Y8;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Type I inner root sheath specific keratin 25 irs3.  
GN Name=KRT25C;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Human scalp;  
 RA Langbein L., Rogers M.A., Winter H., Praetzel S., Schweizer J.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AJ564206; CAD91906.1; -.  
 DR HSSP; P08670; 1GK7.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR Pfam; PF00038; Filament; 1.  
 KW Keratin.  
 SQ SEQUENCE 459 AA; 49823 MW; AFBF29C255D6BD4F CRC64;

Query Match 100.0%; Score 30; DB 2; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 256 LLNNMR 261

#### RESULT 14

Q6R649

ID Q6R649 PRELIMINARY; PRT; 460 AA.  
 AC Q6R649;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Type I keratin intermediate filament C29.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yin J., Li J.Q., Zhou H.M.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY510112; AAS00519.1; -.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR InterPro; IPR002957; Keratin\_I.  
 DR Pfam; PF00038; Filament; 1.  
 DR PRINTS; PR01248; TYPE1KERATIN.  
 KW Keratin.  
 SQ SEQUENCE 460 AA; 49961 MW; 66AAB9DC7D07DE85 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 256 LLNNMR 261

RESULT 15

Q6C2Y6

ID Q6C2Y6 PRELIMINARY; PRT; 461 AA.  
 AC Q6C2Y6;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Similar to DEHA0G23474g Debaryomyces hansenii IPF 4877.1.  
 GN ORFNames=YALIOF04059g;  
 OS Yarrowia lipolytica CLIB99.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Dipodascaceae; Yarrowia.  
 OX NCBI\_TaxID=284591;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swennene D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts.";  
 RL Nature 430:35-44 (2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; CR382132; CAG77783.1; -.  
 DR GO; GO:0030693; F:caspase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 SQ SEQUENCE 461 AA; 47550 MW; 7DF9E3001A072619 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLNNMR 6  
 |||||  
 Db 429 LLNNMR 434

Search completed: February 10, 2005, 15:57:38  
 Job time : 45.7746 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 69.7465 Seconds  
(without alignments)  
44.362 Million cell updates/sec

Title: US-10-067-484-10  
Perfect score: 35  
Sequence: 1 VEASAE LR 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	35	100.0	8	5	ABB81977	Abb81977. 30 kDa ra
2	32	91.4	655	7	ABO73654	Abo73654 Pseudomon
3	31	88.6	31	2	AAR94543	Aar94543 Murine mo
4	31	88.6	393	6	ABU45597	Abu45597 Protein e
5	31	88.6	398	6	ABU47456	Abu47456 Protein e
6	31	88.6	434	4	ABB69842	Abb69842 Drosophil
7	31	88.6	458	7	ADE54253	Ade54253 Rat Prote
8	31	88.6	458	7	ADE54247	Ade54247 Rat Prote
9	31	88.6	458	7	ADE54250	Ade54250 Rat Prote



10	31	88.6	458	7	ADE54244	Ade54244	Rat Prote
11	31	88.6	699	7	ADJ68628	Adj68628	Human hea
12	31	88.6	699	8	ADQ18833	Adq18833	Human sof
13	31	88.6	699	8	ADQ39188	Adq39188	Human myo
14	31	88.6	699	8	ADQ39189	Adq39189	Human myo
15	30	85.7	460	6	ABU22803	Abu22803	Protein e
16	30	85.7	504	7	ABO67399	Abo67399	Klebsiell
17	29	82.9	108	4	AAG91689	Aag91689	C glutami
18	29	82.9	212	5	AAU12144	Aau12144	Brassica
19	29	82.9	213	5	AAU12120	Aau12120	Arabidops
20	29	82.9	215	5	AAU12143	Aau12143	Brassica
21	29	82.9	216	5	AAU12142	Aau12142	Brassica
22	29	82.9	235	7	ABO22380	Abo22380	Mosquito
23	29	82.9	273	6	ABU22950	Abu22950	Protein e
24	29	82.9	359	3	AAV32298	Aay32298	Wheat pol
25	29	82.9	377	7	ABO83966	Abo83966	Pseudomon
26	29	82.9	452	5	ABG91622	Abg91622	Purine/py
27	29	82.9	452	7	ADE62589	Ade62589	Rat Prote
28	29	82.9	452	7	ADD48035	Add48035	Rat Prote
29	29	82.9	452	7	ADE62597	Ade62597	Rat Prote
30	29	82.9	452	7	ADD48031	Add48031	Rat Prote
31	29	82.9	452	7	ADE62593	Ade62593	Rat Prote
32	29	82.9	452	7	ADD48039	Add48039	Rat Prote
33	29	82.9	615	8	ADS21186	Ads21186	Bacterial
34	29	82.9	831	5	ABG91591	Abg91591	Purine/py
35	29	82.9	1165	8	ADN25307	Adn25307	Bacterial
36	28	80.0	75	4	AAM21038	Aam21038	Peptide #
37	28	80.0	75	4	ABB43219	Abb43219	Peptide #
38	28	80.0	75	4	AAM37059	Aam37059	Peptide #
39	28	80.0	75	4	ABB26302	Abb26302	Protein #
40	28	80.0	75	4	AAM76951	Aam76951	Human bon-
41	28	80.0	75	4	AAM64126	Aam64126	Human bra
42	28	80.0	75	4	ABG58613	Abg58613	Human liv
43	28	80.0	75	5	ABG46058	Abg46058	Human pep
44	28	80.0	173	4	AAU06114	Aau06114	Novel hum
45	28	80.0	358	7	ABO82168	Abo82168	Pseudomon

#### ALIGNMENTS

##### RESULT 1

ABB81977

ID ABB81977 standard; peptide; 8 AA.

XX

AC ABB81977;

XX

DT 25-NOV-2002 (first entry)

XX

DE 30. kDa ragweed pollen allergen tryptic peptide 10.

XX

KW Ragweed; pollen; allergen; Ambt 7; glycoprotein; antiallergic;  
KW immunotherapy; disulphide protein.

XX

OS Ambrosia elatior.

XX

PN WO200263012-A2.

XX  
PD 15-AUG-2002.  
XX  
PF 04-FEB-2002; 2002WO-US003346.  
XX  
PR 05-FEB-2001; 2001US-0266686P.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Buchanan BB, Del Val G, Frick OL;  
XX  
DR WPI; 2002-657539/70.  
XX  
PT New ragweed pollen allergens, useful in allergy testing and immunotherapy  
PT regimens, particularly for treating sensitivity to pollen or pollen  
PT allergy (e.g. anaphylaxis, or symptoms of hives or asthma) in a mammal,  
PT especially a human.  
XX  
PS Claim 1; Page 53; 70pp; English.  
XX  
CC The invention relates to an isolated pollen allergen purified from  
CC ragweed pollen, substantially free of any other pollen proteins, or a  
CC protein that is an antigenic fragment of a pollen allergen Ambt 7. The  
CC allergen is characterized by the following physiochemical and biological  
CC properties: (a) being contained in pollen extracts; (b) a glycoprotein;  
CC (c) a sulphhydryl group containing protein; (d) a molecular weight of  
CC about 30 kDa as determined by SDS-polyacrylamide gel electrophoresis; and  
CC (e) possessing allergen activity. The pollen allergen, or antigenic  
CC protein fragment of the pollen allergen Ambt 7, or composition is useful  
CC for treating sensitivity to pollen or pollen allergy in a mammal. This  
CC allergy includes anaphylaxis or atopy, which includes the symptoms of hay  
CC fever, asthma or hives. The allergen is also useful in allergy testing  
CC and immunotherapy regimens. Sequences ABB81968-978 represent tryptic  
CC peptide fragments of the 30 kDa ragweed complete pollen extract  
CC disulphide protein allergen  
XX  
SQ Sequence 8 AA;

Query Match 100.0%; Score 35; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
|||||||  
Db 1 VEASAE LR 8

RESULT 2  
ABO73654  
ID ABO73654 standard; protein; 655 AA.  
XX  
AC ABO73654;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #5829.  
XX

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
 XX  
 OS Pseudomonas aeruginosa.  
 XX  
 PN US6551795-B1.  
 XX  
 PD 22-APR-2003.  
 XX  
 PF 18-FEB-1999; 99US-00252991.  
 XX  
 PR 18-FEB-1998; 98US-0074788P.  
 PR 27-JUL-1998; 98US-0094190P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 XX  
 DR WPI; 2003-615309/58.  
 DR N-PSDB; ABD07225.  
 XX  
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX  
 PS Disclosure; SEQ ID NO 22400; 455pp; English.  
 XX  
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biochip technology. Sequences AB067826-  
 CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX  
 SQ Sequence 655 AA;

Query Match 91.4%; Score 32; DB 7; Length 655;  
 Best Local Similarity 87.5%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 |||||:|  
 Db 602 VEASAD LR 609

RESULT 3  
 AAR94543  
 ID AAR94543 standard; peptide; 31 AA.

XX  
 AC AAR94543;  
 XX  
 DT 18-OCT-1996 (first entry)  
 XX  
 DE Murine moesin antigenic peptide, comprising residues 465-495.  
 XX  
 KW Murine; mouse; moesin; antigenic peptide; rheumatism; diagnosis;  
 KW pre-critical; enzyme linked immunosorbant assay; ELISA; auto-antibodies;  
 KW auto-antigens.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO9607914-A1.  
 XX  
 PD 14-MAR-1996.  
 XX  
 PF 07-SEP-1995; 95WO-JP001776.  
 XX  
 PR 08-SEP-1994; 94JP-00239640.  
 XX  
 PA (FARH ) HOECHST JAPAN LTD.  
 XX  
 PI Wagatsuma M, Kimura M, Watanabe H, Takeuchi F;  
 XX  
 DR WPI; 1996-171727/17.  
 XX  
 PT Detection of auto:antibodies in serum of rheumatism patients - using  
 PT ezrin, radoxin or moesin antigen, for early diagnosis.  
 XX  
 PS Example 6; Page 18; 33pp; Japanese.  
 XX  
 CC The present sequence is an antigenic peptide from murine moesin, which  
 CC can be used to detect auto-antibodies derived from the serum of rheumatic  
 CC patients. For example 250-500 ng of the peptide was coated at 4 degrees C  
 CC onto a microtitre plate, which was then washed and treated with a maltose  
 CC modified casein soln. . After 2hr. at room temp. the plate was dried,  
 CC human serum (diluted 400-fold with casein A soln. contg. 100 microg/ml E.  
 CC coli lysate) added and then incubated at 25 degrees C for 1 hr. . The  
 CC wells were washed, and peroxidase labelled anti-human IgG added. After  
 CC 1hr. at 25 degrees C the wells were washed, developed with  
 CC tetramethylbenzidine and the absorbance at 450 nm measured. The OD(450)  
 CC for a rheumatic patient was slightly greater than that of a healthy  
 CC person. This method enables rheumatism to be diagnosed at an early, pre-  
 CC critical stage of the disease  
 XX  
 SQ Sequence 31 AA;

Query Match 88.6%; Score 31; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
 |||||  
 Db 25 EASAE LR 31

RESULT 4

ABU45597

ID ABU45597 standard; protein; 393 AA.

XX

AC ABU45597;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #31124.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Salmonella paratyphi.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI; 2003-029926/02.

DR N-PSDB; ACA49467.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 25; SEQ ID NO 73521; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 393 AA;

Query Match 88.6%; Score 31; DB 6; Length 393;  
Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
||| |||  
Db 288 VEAGAE LR 295

RESULT 5-

ABU47456

ID ABU47456 standard; protein; 398 AA.

XX

AC ABU47456;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #32983.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS *Salmonella typhi*.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA51326.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 75380; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 398 AA;

Query Match 88.6%; Score 31; DB 6; Length 398;  
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
 ||| |||  
 Db 292 VEAGAE LR 299

RESULT 6  
 ABB69842  
 ID ABB69842 standard; protein; 434 AA.  
 XX  
 AC ABB69842;

XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 36318.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 FA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL13945.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 P3 Disclosure; SEQ ID NO 36318; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 434 AA;

Query Match 88.6%; Score 31; DB 4; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAELR 8  
 |||||  
 Db 284 EASAELR 290

RESULT 7  
 ADE54253  
 ID ADE54253 standard; protein; 458 AA.



XX  
 AC ADE54253;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein BAA24351, SEQ ID NO 56.  
 XX  
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI; 2003-268312/26;  
 DR GENBANK; BAA24351.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 458 AA;

Query Match 88.6%; Score 31; DB 7; Length 458;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||  
Db 298 EASAE LR 304

# RESULT 8

ADE54247

ID ADE54247 standard; protein; 458 AA.

XX

AC ADE54247;

XX

DT 29-JAN-2004 (first entry)

XX

DE Rat Protein BAA24351, SEQ ID NO.50.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

DR GENBANK; BAA24351.

XX

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 458 AA;

Query Match 88.6%; Score 31; DB 7; Length 458;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||  
Db 298 EASAE LR 304

#### RESULT 9

ADE54250

ID ADE54250 standard; protein; 458 AA.

XX

AC ADE54250;

XX

DT 29-JAN-2004 (first entry)

XX

DE Rat Protein BAA24351, SEQ ID NO 53.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI; 2003-268312/26.  
 DR GENBANK; BAA24351.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 458 AA;

Query Match 88.6%; Score 31; DB 7; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAE LR 8  
 |||||  
 Db 298 EASAE LR 304

RESULT 10

ADE54244

ID ADE54244 standard; protein; 458 AA.

XX

AC ADE54244;

XX

DT 29-JAN-2004 (first entry)

XX

DE Rat Protein BAA24351, SEQ ID NO 47.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX

PI Wcolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

DR GENBANK; BAA24351.

XX

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 458 AA;

Query Match 88.6%; Score 31; DB 7; Length 458;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
| | | | |  
Db 298 EASAE LR 304

# RESULT 11

ADJ68628

ID ADJ68628 standard; protein; 699 AA.

XX

AC ADJ68628;

XX

DT 06-MAY-2004 (first entry)

XX

DE Human heat mitochondrial protein as a therapeutic target SeqID434.

XX

KW mitochondrial; human; screening assay; diabetes mellitus;

KW Huntington's disease; osteoarthritis;

KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;

KW osteopathic; ophthalmological; cytostatic.

XX

OS Homo sapiens.

XX

PN WO2003087768-A2.

XX

PD 23-OCT-2003.

XX

PF 04-APR-2003; 2003WO-US010870.

XX

PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389987P.

PR 20-SEP-2002; 2002US-0412418P.

XX

PA (MITO-) MITOKOR.

PA (BUCK-) BUCK INST AGE RES.

XX

PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

PI Warnock DE;

XX

DR WPI; 2003-845369/78.

XX  
PT Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.  
XX  
PS Claim 1; SEQ ID NO 434; 180pp; English.  
XX  
CC This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, nootropic, antidiabetic,  
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
CC cytostatic activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.  
XX  
SQ Sequence 699 AA;

Query Match 88.6%; Score 31; DB 7; Length 699;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||||  
Db 52 EASAE LR 58

RESULT 12  
ADQ18833

ID ADQ18833 standard; protein; 699 AA.  
XX  
AC ADQ18833;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 1652.  
XX  
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2004048938-A2.  
XX  
PD 10-JUN-2004.  
XX  
PF 26-NOV-2003; 2003WO-US038193.  
XX  
PR 26-NOV-2002; 2002US-0429739P.  
XX

PA (PROT-) PROTEIN DESIGN LABS INC.

XX

PI Aziz N, Ginsburg WM, Zlotnik A;

XX

DR WPI; 2004-441208/41.

XX

PT Early detection of soft tissue sarcoma comprises determining expression  
PT of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.

XX

PS Example 2; SEQ ID NO 1652; 210pp; English.

XX

CC The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytostatic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC protein of the invention. The current sequence is not shown within the  
CC specification per se but was submitted in CD format by the inventor.

XX

SQ Sequence 699 AA;

Query Match 88.6%; Score 31; DB 8; Length 699;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAE LR 8

|||||

Db 52 EASAE LR 58

RESULT 13

ADQ39188

ID ADQ39188 standard; protein; 699 AA.

XX

AC ADQ39188;

XX

DT 18-NOV-2004 (first entry)

XX

DE Human myocardial infarction-associated gene derived protein, SEQ ID 851.

XX

KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.

XX

OS Homo sapiens.

XX

PN WO2004058052-A2.

XX

PD 15-JUL-2004.

XX



PF 22-DEC-2003; 2003WO-US040978.  
 XX  
 PR 20-DEC-2002; 2002US-0434778P.  
 PR 10-MAR-2003; 2003US-0453135P.  
 PR 30-APR-2003; 2003US-0466412P.  
 PR 23-SEP-2003; 2003US-0504955P.  
 XX  
 PA (APPL-) APPLERA CORP.  
 XX  
 PI Cargill M, Devlin JJ, Iakoubova O;  
 XX  
 DR WPI; 2004-533949/51.  
 DR N-PSDB; ADQ38360.  
 XX  
 PT Identifying an individual who has an altered risk for developing  
 PT myocardial infarction by detecting a single nucleotide polymorphism in  
 PT the individual's nucleic acids.  
 XX  
 PS Claim 10; SEQ ID NO 851; 145pp; English.  
 XX  
 CC The invention relates to a novel method for identifying an individual who  
 CC has an altered risk for developing myocardial infarction. The method  
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
 CC the nucleotide sequences given in the specification in the individual's  
 CC nucleic acids, where the presence of the SNP is correlated with an  
 CC altered risk for myocardial infarction in the individual. The invention  
 CC further comprises: an isolated nucleic acid molecule comprising at least  
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
 CC the specification or its complement and encoding any one of the amino  
 CC acid sequences given in the specification; an isolated polypeptide  
 CC comprising an amino acid sequence given in the specification; an antibody  
 CC that specifically binds to the polypeptide or its antigen-binding  
 CC fragment; an amplified polynucleotide containing an SNP given in the  
 CC specification and which is between about 16 and 1000 nucleotides in  
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
 CC method for identifying an agent useful in treating or preventing  
 CC myocardial infarction. The novel detection method has cardiact activity.  
 CC The nucleic acids of the invention may be used in gene therapy. The  
 CC method is useful in identifying an individual who has an increased or  
 CC decreased risk for developing myocardial infarction and for preparing a  
 CC composition for treating or preventing myocardial infarction. This  
 CC sequence represents the protein of a human myocardial infarction-  
 CC associated gene containing one or more SNP's of the invention. Note: This  
 CC sequence was not shown in the specification. The sequence has come from  
 CC an electronic sequence listing downloaded from the WIPO website.  
 XX  
 SQ Sequence 699 AA;

Query Match 88.6%; Score 31; DB 8; Length 699;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAE LR 8  
 |||||  
 Db 52 EASAE LR 58

RESULT 14

ADQ39189

ID ADQ39189 standard; protein; 699 AA.

XX

AC ADQ39189;

XX

DT 18-NOV-2004 (first entry)

XX

DE Human myocardial infarction-associated gene derived protein, SEQ ID 852.

XX

KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.

XX

OS Homo sapiens.

XX

PN WO2004058052-A2.

XX

PD 15-JUL-2004.

XX

PF 22-DEC-2003; 2003WO-US040978.

XX

PR 20-DEC-2002; 2002US-0434778P.

PR 10-MAR-2003; 2003US-0453135P.

PR 30-APR-2003; 2003US-0466412P.

PR 23-SEP-2003; 2003US-0504955P.

XX

PA (APPL-) APPLERA CORP.

XX

PI Cargill M, Devlin JJ, Iakoubova O;

XX

DR WPI; 2004-533949/51.

DR N-PSDB; ADQ38361.

XX

PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.

XX

PS Claim 10; SEQ ID NO 852; 145pp; English.

XX

CC The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a

CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC sequence represents the protein of a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.  
XX  
SQ Sequence 699 AA;

Query Match 88.6%; Score 31; DB 8; Length 699;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
|||  
Db 52 EASAELR 58

RESULT 15

ABU22803

ID ABU22803 standard; protein; 460 AA.

XX

AC ABU22803;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #8330.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Burkholderia mallei.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI; 2003-029926/02.

DR N-PSDB; ACA26673.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 25; SEQ ID NO 50727; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 460 AA;

Query Match 85.7%; Score 30; DB 6; Length 460;  
Best Local Similarity 87.5%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
|| |||||  
Db 207 VELSAE LR 214

Search completed: February 10, 2005, 15:48:47  
Job time : 70.7465 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 17.9155 Seconds  
 (without alignments)  
 33.334 Million cell updates/sec

Title: US-10-067-484-10  
 Perfect score: 35  
 Sequence: 1 VEASAE LR 8

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	32	91.4	655	4	US-09-252-991A-22400	Sequence 22400, A
2	31	88.6	31	2	US-08-793-937-3	Sequence 3, Appli
3	31	88.6	31	3	US-09-235-283-3	Sequence 3, Appli
4	31	88.6	699	4	US-09-538-092-995	Sequence 995, App
5	30	85.7	504	4	US-09-489-039A-13916	Sequence 13916, A
6	29	82.9	235	4	US-09-345-236B-13	Sequence 13, Appl
7	29	82.9	359	4	US-09-699-266A-11	Sequence 11, Appl
8	29	82.9	377	4	US-09-252-991A-32712	Sequence 32712, A
9	29	82.9	395	4	US-09-902-540-14097	Sequence 14097, A
10	29	82.9	411	4	US-09-949-016-6312	Sequence 6312, Ap
11	29	82.9	414	4	US-09-949-016-9493	Sequence 9493, Ap
12	29	82.9	452	4	US-09-563-794B-140	Sequence 140, App
13	28	80.0	358	4	US-09-252-991A-30914	Sequence 30914, A
14	28	80.0	468	4	US-09-902-540-13224	Sequence 13224, A
15	28	80.0	583	4	US-09-902-540-14031	Sequence 14031, A
16	28	80.0	686	4	US-09-252-991A-26437	Sequence 26437, A
17	28	80.0	942	4	US-09-489-039A-13982	Sequence 13982, A

18	27	77.1	92	4	US-09-640-211A-2185	Sequence 2185, Ap
19	27	77.1	186	4	US-09-252-991A-16962	Sequence 16962, A
20	27	77.1	198	4	US-09-640-211A-2193	Sequence 2193, Ap
21	27	77.1	205	4	US-09-949-016-8356	Sequence 8356, Ap
22	27	77.1	210	4	US-09-053-374A-7	Sequence 7, Appli
23	27	77.1	218	1	US-08-247-946A-3	Sequence 3, Appli
24	27	77.1	218	5	PCT-US95-06420-3	Sequence 3, Appli
25	27	77.1	310	4	US-09-252-991A-22740	Sequence 22740, A
26	27	77.1	320	4	US-09-252-991A-32032	Sequence 32032, A
27	27	77.1	370	4	US-09-328-352-5826	Sequence 5826, Ap
28	27	77.1	373	4	US-09-252-991A-28902	Sequence 28902, A
29	27	77.1	415	4	US-09-107-532A-6504	Sequence 6504, Ap
30	27	77.1	462	3	US-09-036-987A-18	Sequence 18, Appl
31	27	77.1	462	3	US-09-370-700-18	Sequence 18, Appl
32	27	77.1	462	4	US-09-603-207-18	Sequence 18, Appl
33	27	77.1	508	4	US-09-252-991A-24250	Sequence 24250, A
34	27	77.1	508	4	US-09-252-991A-32651	Sequence 32651, A
35	27	77.1	565	4	US-09-252-991A-27022	Sequence 27022, A
36	27	77.1	611	4	US-09-252-991A-20097	Sequence 20097, A
37	27	77.1	654	4	US-09-949-016-8841	Sequence 8841, Ap
38	27	77.1	654	4	US-09-949-016-8842	Sequence 8842, Ap
39	27	77.1	1297	3	US-09-540-245A-17	Sequence 17, Appl
40	27	77.1	1420	4	US-09-902-540-11946	Sequence 11946, A
41	27	77.1	1554	4	US-09-252-991A-26814	Sequence 26814, A
42	27	77.1	1798	4	US-09-845-583A-8	Sequence 8, Appli
43	27	77.1	1798	4	US-09-561-709B-11	Sequence 11, Appl
44	27	77.1	1798	4	US-09-917-254-87	Sequence 87, Appl
45	27	77.1	1854	4	US-09-949-016-11625	Sequence 11625, A

#### ALIGNMENTS

##### RESULT 1

US-09-252-991A-22400

; Sequence 22400, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22400

; LENGTH: 655

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22400

Query Match

91.4%; Score 32; DB 4; Length 655;

Best Local Similarity 87.5%; Pred. No. 74;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
|||:|  
Db 602 VEASADLR 609

RESULT 2

US-08-793-937-3

; Sequence 3, Application US/08793937  
; Patent No. 5976810  
; GENERAL INFORMATION:  
; APPLICANT: WAGATSUMA, Masako  
; APPLICANT: KIMURA, Michio  
; APPLICANT: WATANABE, Hiroshi  
; APPLICANT: TAKEUCHI, Fujio  
; TITLE OF INVENTION: METHOD OF DETECTING AUTOANTIBODY PRESENT IN THE SERUM  
; TITLE OF INVENTION: OF RHEUMATIC  
; FILE REFERENCE: 05552.1414-00000  
; CURRENT APPLICATION NUMBER: US/08/793,937  
; CURRENT FILING DATE: 1997-05-08  
; EARLIER APPLICATION NUMBER: PCT/JP95/01776  
; EARLIER FILING DATE: 1995-09-07  
; EARLIER APPLICATION NUMBER: JAPAN 6/239640  
; EARLIER FILING DATE: 1994-09-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-08-793-937-3

Query Match 88.6%; Score 31; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||  
Db 25 EASAE LR 31

RESULT 3

US-09-235-283-3

; Sequence 3, Application US/09235283  
; Patent No. 6225442  
; GENERAL INFORMATION:  
; APPLICANT: WAGATSUMA, Masako  
; APPLICANT: KIMURA, Michio  
; APPLICANT: WATANABE, Hiroshi  
; APPLICANT: TAKEUCHI, Fujio  
; TITLE OF INVENTION: METHOD OF DETECTING AUTOANTIBODY PRESENT IN THE SERUM  
; TITLE OF INVENTION: OF RHEUMATIC  
; FILE REFERENCE: 05552.1414-00000  
; CURRENT APPLICATION NUMBER: US/09/235,283  
; CURRENT FILING DATE: 1999-01-22

; PRIOR APPLICATION NUMBER: 08/793,937  
; PRIOR FILING DATE: 1997-05-08  
; PRIOR APPLICATION NUMBER: JAPAN 6/239640  
; PRIOR FILING DATE: 1994-09-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-235-283-3

Query Match 88.6%; Score 31; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||  
Db 25 EASAE LR 31

RESULT 4

US-09-538-092-995  
; Sequence 995, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqFormatter Version 0.9  
; SEQ ID NO 995  
; LENGTH: 699  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number P23327  
US-09-538-092-995

Query Match 88.6%; Score 31; DB 4; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||  
Db 52 EASAE LR 58



RESULT 5

US-09-489-039A-13916

; Sequence 13916, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; CURRENT FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 13916  
 ; LENGTH: 504  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-13916

Query Match 85.7%; Score 30; DB 4; Length 504;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1.VEASAE LR 8  
 ||||::|  
 Db 318 VEASAQIR 325

RESULT 6

US-09-345-236B-13

; Sequence 13, Application US/09345236B  
 ; Patent No. 6521454  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Becnel, James J.  
 ; APPLICANT: Tuku, Fukuda  
 ; APPLICANT: Moser, Bettina  
 ; APPLICANT: Cockburn, Andrew  
 ; APPLICANT: White, Susan E.  
 ; APPLICANT: Undeen, Albert H.  
 ; TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal  
 ; TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates  
 ; FILE REFERENCE: 21042.0004  
 ; CURRENT APPLICATION NUMBER: US/09/345,236B  
 ; CURRENT FILING DATE: 1999-06-30  
 ; NUMBER OF SEQ ID NOS: 148  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 13  
 ; LENGTH: 235  
 ; TYPE: PRT  
 ; ORGANISM: mosquito baculovirus  
 US-09-345-236B-13

Query Match 82.9%; Score 29; DB 4; Length 235;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy            1 VEASAE LR 8  
              |:|:|:|:|  
Db            90 VDSSAE LR 97

RESULT 7

US-09-699-266A-11  
; Sequence 11, Application US/09699266A  
; Patent No. 6559354  
; GENERAL INFORMATION:  
; APPLICANT: Ma, Hongchang  
; APPLICANT: Morakinyo, Layo O.  
; APPLICANT: Odell, Joan T.  
; APPLICANT: Orozco Jr., Emil M.  
; APPLICANT: Rafalski, J. Antoni  
; TITLE OF INVENTION: TRANSCRIPTION AND GENE EXPRESSION REGULATORS  
; FILE REFERENCE: BB1164 US NA  
; CURRENT APPLICATION NUMBER: US/09/699,266A  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: PCT/US99/08385  
; PRIOR FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/083,212  
; PRIOR FILING DATE: 1998-04-27  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 11  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-699-266A-11

Query Match            82.9%; Score 29; DB 4; Length 359;  
Best Local Similarity   85.7%; Pred. No. 1.7e+02;  
Matches        6; Conservative    1; Mismatches    0; Indels        0; Gaps        0;

Qy            1 VEASAE L 7  
              :|:|:|:|  
Db            318 IEASAE L 324

RESULT 8

US-09-252-991A-32712  
; Sequence 32712, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32712  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32712

Query Match 82.9%; Score 29; DB 4; Length 377;  
Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
|||:|:|  
Db 193 VEATAQLR 200

RESULT 9

US-09-902-540-14097  
; Sequence 14097, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 14097  
; LENGTH: 395  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-14097

Query Match 82.9%; Score 29; DB 4; Length 395;  
Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
:|:|:|  
Db 247 LEAAAE LR 254

RESULT 10

US-09-949-016-6312  
; Sequence 6312, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6312  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6312

Query Match 82.9%; Score 29; DB 4; Length 411;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
: ||: |||  
Db 94 LEAAAE LR 101

RESULT 11

US-09-949-016-9493  
; Sequence 9493, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES  
THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9493  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9493

Query Match 82.9%; Score 29; DB 4; Length 414;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
: ||: |||

Db 97 LEAAAELR 104

RESULT 12

US-09-563-794B-140  
; Sequence 140, Application US/09563794B  
; Patent No. 6808876  
; GENERAL INFORMATION:  
; APPLICANT: KRUGER, MARTIN  
; APPLICANT: WELCH, PETER J.  
; APPLICANT: BARBER, JACK R.  
; TITLE OF INVENTION: CELLULAR REGULATORS OF INFECTIOUS AGENTS AND METHODS OF  
; TITLE OF INVENTION: USE  
; FILE REFERENCE: 039316-0801  
; CURRENT APPLICATION NUMBER: US/09/563,794B  
; CURRENT FILING DATE: 2000-05-02  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 140  
; LENGTH: 452  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-563-794B-140

Query Match 82.9%; Score 29; DB 4; Length 452;  
Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAELR 8  
||| ||:|  
Db 415 VEAGAEIR 422

RESULT 13

US-09-252-991A-30914  
; Sequence 30914, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30914  
; LENGTH: 358  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30914

Query Match 80.0%; Score 28; DB 4; Length 358;

Best Local Similarity 75.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
:|||||  
Db 291 LEASAEFR 298

RESULT 14

US-09-902-540-13224

; Sequence 13224, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 13224  
; LENGTH: 468  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus

US-09-902-540-13224

Query Match 80.0%; Score 28; DB 4; Length 468;  
Best Local Similarity 75.0%; Pred. No. 3.8e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
|||||:  
Db 100 VEASQEMR 107

RESULT 15

US-09-902-540-14031

; Sequence 14031, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 14031  
; LENGTH: 583

; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-14031

Query Match 80.0%; Score 28; DB 4; Length 583;  
Best Local Similarity 85.7%; Pred. No. 4.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
||:||||  
Db 170 EATAELR 176

Search completed: February 10, 2005, 16:02:11  
Job time : 18.9155 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:49:10 ; Search time 47.8873 Seconds  
(without alignments)  
54.586 Million cell updates/sec

Title: US-10-067-484-10  
Perfect score: 35  
Sequence: 1 VEASAE LR 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	% Query		Match Length	DB	ID	Description
	Score	Match				
1	35	100.0	8	14	US-10-067-484-10	Sequence 10, Appl
2	35	100.0	8	14	US-10-067-620-10	Sequence 10, Appl
3	31	88.6	393	15	US-10-282-122A-73521	Sequence 73521, A
4	31	88.6	398	15	US-10-282-122A-75380	Sequence 75380, A
5	31	88.6	699	16	US-10-408-765A-434	Sequence 434, App
6	31	88.6	699	17	US-10-741-600-851	Sequence 851, App
7	31	88.6	699	17	US-10-741-600-852	Sequence 852, App
8	30	85.7	460	15	US-10-282-122A-50727	Sequence 50727, A
9	29	82.9	108	9	US-09-738-626-5443	Sequence 5443, Ap
10	29	82.9	210	14	US-10-156-761-13395	Sequence 13395, A
11	29	82.9	212	15	US-10-312-222-39	Sequence 39, Appl
12	29	82.9	213	15	US-10-312-222-4	Sequence 4, Appli
13	29	82.9	215	15	US-10-312-222-36	Sequence 36, Appl
14	29	82.9	216	15	US-10-312-222-33	Sequence 33, Appl
15	29	82.9	273	15	US-10-282-122A-50874	Sequence 50874, A
16	29	82.9	464	15	US-10-424-599-229030	Sequence 229030,
17	29	82.9	530	14	US-10-156-761-10819	Sequence 10819, A
18	29	82.9	615	15	US-10-369-493-10219	Sequence 10219, A
19	29	82.9	1165	15	US-10-369-493-7960	Sequence 7960, Ap
20	28	80.0	75	9	US-09-864-761-41600	Sequence 41600, A
21	28	80.0	104	16	US-10-767-701-57029	Sequence 57029, A
22	28	80.0	105	15	US-10-424-599-151957	Sequence 151957,
23	28	80.0	120	15	US-10-424-599-165503	Sequence 165503,
24	28	80.0	146	15	US-10-425-114-44124	Sequence 44124, A
25	28	80.0	165	16	US-10-767-701-46380	Sequence 46380, A
26	28	80.0	173	9	US-09-779-307-8	Sequence 8, Appli
27	28	80.0	173	9	US-09-779-307-24	Sequence 24, Appl
28	28	80.0	178	16	US-10-767-701-39377	Sequence 39377, A
29	28	80.0	220	16	US-10-767-701-34576	Sequence 34576, A
30	28	80.0	266	15	US-10-424-599-262111	Sequence 262111,
31	28	80.0	270	15	US-10-425-114-45812	Sequence 45812, A
32	28	80.0	294	16	US-10-767-701-46379	Sequence 46379, A
33	28	80.0	371	9	US-09-738-626-6691	Sequence 6691, Ap
34	28	80.0	398	9	US-09-815-242-10390	Sequence 10390, A
35	28	80.0	398	15	US-10-369-493-23607	Sequence 23607, A
36	28	80.0	398	15	US-10-282-122A-55759	Sequence 55759, A
37	28	80.0	398	15	US-10-282-122A-56745	Sequence 56745, A
38	28	80.0	400	15	US-10-424-599-165501	Sequence 165501,
39	28	80.0	413	14	US-10-156-761-12219	Sequence 12219, A



40	28	80.0	550	15	US-10-282-122A-53663	Sequence 53663, A
41	28	80.0	577	15	US-10-236-031B-54	Sequence 54, Appl
42	28	80.0	577	15	US-10-360-849A-27	Sequence 27, Appl
43	28	80.0	577	16	US-10-408-765A-453	Sequence 453, App
44	28	80.0	577	16	US-10-408-765A-454	Sequence 454, App
45	28	80.0	577	16	US-10-648-593-187	Sequence 187, App

# ALIGNMENTS

## RESULT 1

US-10-067-484-10

```
; Sequence 10, Application US/10067484
; Publication No. US20030170763A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; TITLE OF INVENTION: RAGWEED ALLERGENS
; FILE REFERENCE: 416272000200
; CURRENT APPLICATION NUMBER: US/10/067,484
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Ragweed
```

US-10-067-484-10

```
Query Match          100.0%; Score 35; DB 14; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 VEASAE LR 8
        |||||
Db      1 VEASAE LR 8
```

## RESULT 2

US-10-067-620-10

```
; Sequence 10, Application US/10067620
; Publication No. US20030180225A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; APPLICANT: Teuber, Suzanne S.
; TITLE OF INVENTION: WALNUT AND RYEGRASS ALLERGENS
; FILE REFERENCE: 416272003400
; CURRENT APPLICATION NUMBER: US/10/067,620
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
```

; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Ragweed  
US-10-067-620-10

Query Match 100.0%; Score 35; DB 14; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
|||  
Db 1 VEASAE LR 8

RESULT 3

US-10-282-122A-73521  
; Sequence 73521, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 73521  
; LENGTH: 393  
; TYPE: PRT  
; ORGANISM: Salmonella paratyphi A  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (373)..(373)  
; OTHER INFORMATION: X=any amino acid  
US-10-282-122A-73521

Query Match 88.6%; Score 31; DB 15; Length 393;  
Best Local Similarity 87.5%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| |||  
Db 288 VEAGAE LR 295

RESULT 4

US-10-282-122A-75380  
; Sequence 75380, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 75380  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Salmonella typhi  
US-10-282-122A-75380

Query Match 88.6%; Score 31; DB 15; Length 398;  
Best Local Similarity 87.5%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| |||  
Db 292 VEAGAE LR 299

RESULT 5

US-10-408-765A-434

; Sequence 434, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 434  
; LENGTH: 699  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-434

Query Match 88.6%; Score 31; DB 16; Length 699;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
|||||  
Db 52 EASAE LR 58

RESULT 6

US-10-741-600-851

; Sequence 851, Application US/10741600

; Publication No. US20050026169A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES  
THEREOF

; FILE REFERENCE: CL001499

; CURRENT APPLICATION NUMBER: US/10/741,600

; CURRENT FILING DATE: 2003-12-22

; NUMBER OF SEQ ID NOS: 73997

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 851

; LENGTH: 699

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-741-600-851

Query Match 88.6%; Score 31; DB 17; Length 699;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8

|||||

Db 52 EASAE LR 58

RESULT 7

US-10-741-600-852

; Sequence 852, Application US/10741600

; Publication No. US20050026169A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES  
THEREOF

; FILE REFERENCE: CL001499

; CURRENT APPLICATION NUMBER: US/10/741,600

; CURRENT FILING DATE: 2003-12-22

; NUMBER OF SEQ ID NOS: 73997

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 852

; LENGTH: 699

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-741-600-852

Query Match 88.6%; Score 31; DB 17; Length 699;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8

|||||

## RESULT 8

US-10-282-122A-50727

; Sequence 50727, Application US/10282122A

; Publication No. US20040029129A1

## ; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 50727

; LENGTH: 460

; TYPE: PRT

; ORGANISM: Burkholderia mallei

US-10-282-122A-50727

Query Match 85.7%; Score 30; DB 15; Length 460;

Best Local Similarity 87.5%; Pred. No. 4e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
|| ||||  
Db 207 VELSAE LR 214

RESULT 9

US-09-738-626-5443  
; Sequence 5443, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 5443  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5443

Query Match 82.9%; Score 29; DB 9; Length 108;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||:|:|  
Db 89 VESSSEL R 96

RESULT 10

US-10-156-761-13395  
; Sequence 13395, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13395  
; LENGTH: 210  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13395

Query Match 82.9%; Score 29; DB 14; Length 210;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
: ||: ||||  
Db 72 LEAAAE LR 79

RESULT 11

US-10-312-222-39

; Sequence 39, Application US/10312222  
; Publication No. US20040093633A1  
; GENERAL INFORMATION:  
; APPLICANT: Xiao, Shun Y  
; APPLICANT: Turner, John G  
; APPLICANT: Coleman, Mark  
; APPLICANT: Ellwood, Simon  
; TITLE OF INVENTION: Plant Resistance Gene  
; FILE REFERENCE: 0380-PO3051US0  
; CURRENT APPLICATION NUMBER: US/10/312,222  
; CURRENT FILING DATE: 2002-12-25  
; PRIOR APPLICATION NUMBER: GB 0015122.5  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Brassica rapa  
US-10-312-222-39

Query Match 82.9%; Score 29; DB 15; Length 212;  
Best Local Similarity 87.5%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| ||||  
Db 83 VEAYAE LR 90



RESULT 12

US-10-312-222-4

; Sequence 4, Application US/10312222  
 ; Publication No. US20040093633A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xiao, Shun Y  
 ; APPLICANT: Turner, John G  
 ; APPLICANT: Coleman, Mark  
 ; APPLICANT: Ellwood, Simon  
 ; TITLE OF INVENTION: Plant Resistance Gene  
 ; FILE REFERENCE: 0380-PO3051US0  
 ; CURRENT APPLICATION NUMBER: US/10/312,222  
 ; CURRENT FILING DATE: 2002-12-25  
 ; PRIOR APPLICATION NUMBER: GB 0015122.5  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 75  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 213  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 US-10-312-222-4

Query Match 82.9%; Score 29; DB 15; Length 213;  
 Best Local Similarity 87.5%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 ||| |||  
 Db 83 VEAYAE LR 90

RESULT 13

US-10-312-222-36

; Sequence 36, Application US/10312222  
 ; Publication No. US20040093633A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xiao, Shun Y  
 ; APPLICANT: Turner, John G  
 ; APPLICANT: Coleman, Mark  
 ; APPLICANT: Ellwood, Simon  
 ; TITLE OF INVENTION: Plant Resistance Gene  
 ; FILE REFERENCE: 0380-PO3051US0  
 ; CURRENT APPLICATION NUMBER: US/10/312,222  
 ; CURRENT FILING DATE: 2002-12-25  
 ; PRIOR APPLICATION NUMBER: GB 0015122.5  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 75  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 36  
 ; LENGTH: 215  
 ; TYPE: PRT  
 ; ORGANISM: Brassica rapa  
 US-10-312-222-36

Query Match 82.9%; Score 29; DB 15; Length 215;

Best Local Similarity 87.5%; Pred. No. 2.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| |||  
Db 83 VEAYAE LR 90

RESULT 14

US-10-312-222-33

; Sequence 33, Application US/10312222  
; Publication No. US20040093633A1  
; GENERAL INFORMATION:  
; APPLICANT: Xiao, Shun Y  
; APPLICANT: Turner, John G  
; APPLICANT: Coleman, Mark  
; APPLICANT: Ellwood, Simon  
; TITLE OF INVENTION: Plant Resistance Gene  
; FILE REFERENCE: 0380-PO3051US0  
; CURRENT APPLICATION NUMBER: US/10/312,222  
; CURRENT FILING DATE: 2002-12-25  
; PRIOR APPLICATION NUMBER: GB 0015122.5  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Brassica rapa  
US-10-312-222-33

Query Match 82.9%; Score 29; DB 15; Length 216;  
Best Local Similarity 87.5%; Pred. No. 2.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| |||  
Db 83 VEAYAE LR 90

RESULT 15

US-10-282-122A-50874

; Sequence 50874, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 50874  
 ; LENGTH: 273  
 ; TYPE: PRT  
 ; ORGANISM: Bordetella pertussis  
 US-10-282-122A-50874

Query Match 82.9%; Score 29; DB 15; Length 273;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 | : | | | : |  
 Db 22 VDASAEVR 29

Search completed: February 10, 2005, 16:41:34  
 Job time : 47.8873 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 12.3944 Seconds  
 (without alignments)  
 62.104 Million cell updates/sec

Title: US-10-067-484-10  
 Perfect score: 35

Sequence: 1 VEASAE LR 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	32	91.4	656	2	C83479	probable proteinase
2	31	88.6	398	2	AI0973	glycine C-acetyltr
3	31	88.6	699	2	A54660	histidine rich cal
4	30	85.7	304	2	G95859	probable oligopept
5	30	85.7	304	2	AD3168	hypothetical prote
6	30	85.7	532	2	E87343	ABC transporter, A
7	29	82.9	147	2	T36374	probable acetyltra
8	29	82.9	213	2	T46078	hypothetical prote
9	29	82.9	213	2	T30116	hypothetical prote
10	29	82.9	452	2	S72266	translation initia
11	29	82.9	553	2	F91063	DNA repair protein
12	29	82.9	553	2	B85908	protein used in re
13	29	82.9	2962	2	T19756	hypothetical prote
14	28	80.0	233	2	S60208	fomE protein - Str
15	28	80.0	314	2	B83487	hypothetical prote
16	28	80.0	364	2	I40168	carbamoyl-phosphat
17	28	80.0	398	1	XUECGA	glycine C-acetyltr
18	28	80.0	398	2	H86037	glycine C-acetyltr
19	28	80.0	398	2	G91190	2-amino-3-ketobuty
20	28	80.0	427	2	H87285	phosphoribosylamin
21	28	80.0	434	2	AG1697	hypothetical prote
22	28	80.0	440	2	D71715	hypothetical prote
23	28	80.0	545	2	H83079	hypothetical prote
24	28	80.0	571	2	AC1216	ABC transporters,
25	28	80.0	571	2	AH1569	ABC transporters,
26	28	80.0	577	1	A41289	moesin - human

27	28	80.0	577	1	S39804	moesin - pig
28	28	80.0	1466	2	T30566	ATP-binding cassette
29	28	80.0	1687	2	S41742	calcium channel al
30	28	80.0	2314	2	T28698	hypothetical prote
31	28	80.0	7962	2	I38346	elastic titin - hu
32	27	77.1	90	2	A31512	parathymosin - bov
33	27	77.1	102	2	A32264	parathymosin - hum
34	27	77.1	102	2	B31512	parathymosin - rat
35	27	77.1	106	2	S20422	zinc-binding prote
36	27	77.1	218	1	TVHURR	transforming prote
37	27	77.1	219	1	TVBYSR	transforming prote
38	27	77.1	219	1	TVBYPR	ras-like protein 1
39	27	77.1	219	2	T45545	transforming prote
40	27	77.1	224	2	F70048	serine proteinase
41	27	77.1	238	2	C69334	2-hydroxy-6-oxohep
42	27	77.1	245	2	T09569	MADS box protein M
43	27	77.1	276	2	H81931	probable bis(5'-nu
44	27	77.1	291	2	AD1303	conserved hypothet
45	27	77.1	291	2	AD1675	conserved hypothet

# ALIGNMENTS

## RESULT 1

C83479

probable proteinase PA1327 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C;Species: *Pseudomonas aeruginosa*

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C;Accession: C83479

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83479

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-656 <STO>

A;Cross-references: UNIPROT:Q9I417; GB:AE004562; GB:AE004091; NID:g9947263; PIDN:AAG04716.1; GSPDB:GN00131; PASP:PA1327

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA1327

Query Match 91.4%; Score 32; DB 2; Length 656;  
 Best Local Similarity 87.5%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 |||||:|  
 Db 589 VEASAD LR 596

RESULT 2

AI0973

glycine C-acetyltransferase (EC 2.3.1.29) - *Salmonella enterica* subsp. *enterica* serovar Typhi (strain CT18)

C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi

A;Note: this species has also been called *Salmonella typhi*

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AI0973

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mungall, K.L.; Bentley, S.D.; Holden, M.T.G.; Sebaihia, M.; Baker, S.; Basham, D.; Brooks, K.; Chillingworth, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Feltwell, T.; Hamlin, N.; Haque, A.; Hien, T.T.; Holroyd, S.; Jagels, K.; Krogh, A.; Larsen, T.S.; Leather, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.

A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18.

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AI0973

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-398 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD03285.1; PID:g16504906; GSPDB:GN00176

C;Genetics:

A;Gene: STY4086

C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology

C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate

F;244/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 88.6%; Score 31; DB 2; Length 398;  
Best Local Similarity 87.5%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| |||  
Db 292 VEAGAE LR 299

RESULT 3

A54660

histidine rich calcium binding protein - human

C;Species: *Homo sapiens* (man)

C;Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 09-Jul-2004

C;Accession: A54660

R;Hofmann, S.L.; Topham, M.; Hsieh, C.L.; Francke, U.

Genomics 9, 656-669, 1991

A;Title: cDNA and genomic cloning of HRC, a human sarcoplasmic reticulum protein, and localization of the gene to human chromosome 19 and mouse chromosome 7.

A;Reference number: A54660; MUID:91244309; PMID:2037293

A;Accession: A54660

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-699 <HOF>  
A;Cross-references: UNIPROT:P23327; GB:M60052; NID:g183918; PIDN:AAA88071.1;  
PID:g183919  
C;Genetics:  
A;Gene: GDB:HRC  
A;Cross-references: GDB:126369; OMIM:142705  
A;Map position: 19q13.3-19q13.3  
C;Keywords: calcium binding

Query Match 88.6%; Score 31; DB 2; Length 699;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
|||||||  
Db 52 EASAELR 58

RESULT 4

G95859

probable oligopeptide ABC transporter permease protein SMb20143 [imported] -  
Sinorhizobium meliloti (strain 1021) megaplasmid pSymB

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C;Accession: G95859

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernandez-Lucas, I.; Becker, A.; Cowie, A.; Gouzy, J.; Golding, B.; Puhler, A.

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: G95859

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-304 <KUR>

A;Cross-references: UNIPROT:Q92X19; GB:AL591985; PIDN:CAC48543.1; PID:g15140015; GSPDB:GN00167

A;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Barnett, M.J.; Becker, A.; Boistard, P.; Bothe, G.; Boutry, M.; Bowser, L.; Buhrmester, J.; Cadieu, E.; Capela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Gloux, S.; Godrie, T.; Goffeau, A.; Golding, B.; Gouzy, J.; Gurjal, M.; Hernandez-Lucas, I.; Hong, A.; Huizar, L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, V.; Masuy, D.; Palm, C.; Peck, M.C.; Pohl, T.M.; Portetelle, D.; Purnelle, B.; Ramsperger, U.; Surzycki, R.; Thebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.; Batut, J.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMb20143

A;Genome: plasmid

Query Match 85.7%; Score 30; DB 2; Length 304;  
Best Local Similarity 75.0%; Pred. No. 35;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| ||| ::  
Db 40 VEASAE MK 47

RESULT 5

AD3168

hypothetical protein Atu5070 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid AT

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C;Accession: AD3168

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitajima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tingey, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens. C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AD3168

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-304 <KUR>

A;Cross-references: UNIPROT:Q8UKN0; GB:AE008687; PIDN:AAL45762.1; PID:g17743495; GSPDB:GN00188

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu5070

A;Genome: plasmid

Query Match 85.7%; Score 30; DB 2; Length 304;  
Best Local Similarity 75.0%; Pred. No. 35;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| ||| ::  
Db 40 VEASAE MK 47

RESULT 6

E87343

ABC transporter, ATP-binding protein Cydc CC0760 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C;Accession: E87343



R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: E87343

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-532 <STO>

A;Cross-references: UNIPROT:Q9AA46; GB:AE005673; NID:g13421995; PIDN:AAK22745.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC0760

C;Superfamily: Mycobacterium tuberculosis probable ABC transporter cydC; ATP-binding cassette homology

Query Match 85.7%; Score 30; DB 2; Length 532;  
Best Local Similarity 87.5%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
| | | | |  
Db 214 VSASAE LR 221

#### RESULT 7

T36374

probable acetyltransferase - *Streptomyces coelicolor*

C;Species: *Streptomyces coelicolor*

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: T36374

R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.;

Rajandream, M.A.

submitted to the EMBL Data Library, April 1999

A;Reference number: Z21573

A;Accession: T36374

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-147 <OLI>

A;Cross-references: UNIPROT:Q9X8M0; EMBL:AL049628; PIDN:CAB40863.1;

GSPDB:GN00070; SCOEDB:SCE94.14

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SCE94.14

Query Match 82.9%; Score 29; DB 2; Length 147;  
Best Local Similarity 87.5%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
| | | | |  
Db 14 VEAVAE LR 21

RESULT 8

T46078

hypothetical protein T20E23.70 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C;Accession: T46078

R;Bargues, M.; Perez-Perez, A.; Terol, J.; Torres, A.; Perez-Alonso, M.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
submitted to the Protein Sequence Database, December 1999

A;Reference number: Z23020

A;Accession: T46078

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-213 <BAR>

A;Cross-references: UNIPROT:Q9SCS7; EMBL:AL133363

A;Experimental source: cultivar Columbia; BAC clone T20E23

C;Genetics:

A;Map position: 3

A;Introns: 99/2

A;Note: T20E23.70

C;Superfamily: Arabidopsis thaliana hypothetical protein T20E23.60

Query Match 82.9%; Score 29; DB 2; Length 213;  
Best Local Similarity 87.5%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAELR 8  
||| |||  
Db 33 VEAYAE LR 90

RESULT 9

T30116

hypothetical protein F55F1.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C;Accession: T30116

R;Fulton, L.; Gattung, S.

submitted to the EMBL Data Library, May 1996

A;Description: The sequence of C. elegans cosmid F55F1.

A;Reference number: Z20738

A;Accession: T30116

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-213 <FUL>

A;Cross-references: EMBL:U58744; PIDN:AAB00619.1; GSPDB:GN00028; CESP:F55F1.1

A;Experimental source: strain Bristol N2; clone F55F1

C;Genetics:

A;Gene: CESP:F55F1.1

A;Map position: X

A;Introns: 3/3; 40/3; 79/3; 129/3

Query Match 82.9%; Score 29; DB 2; Length 213;  
Best Local Similarity 75.0%; Pred. No. 42;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
:|:|||||  
Db 44 LESSAE LR 51

RESULT 10

S72266

translation initiation factor eIF2B gamma chain - rat

N;Alternate names: guanine nucleotide-exchange protein

C;Species: Rattus norvegicus (Norway rat)

C;Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 09-Jul-2004

C;Accession: S72266

R;Price, N.T.; Kimball, S.R.; Jefferson, L.S.; Proud, C.G.

Biochem. J. 318, 631-636, 1996

A;Title: Cloning of cDNA for the gamma-subunit of mammalian translation initiation factor 2B, the guanine nucleotide-exchange factor for eukaryotic initiation factor 2.

A;Reference number: S72265; MUID:96404916; PMID:8809057

A;Accession: S72266

A;Molecule type: mRNA

A;Residues: 1-452 <PRI>

A;Cross-references: UNIPROT:P70541; EMBL:U38253; NID:g1537014; PIDN:AAC52788.1; PID:g1537015

A;Experimental source: strain Sprague-Dawley

C;Complex: heteropentamer; alpha, beta, gamma, delta and epsilon chain

C;Function:

A;Description: responsible for recycling initiation factor eIF2 to its active GTP-bound state; binds pyridine dinucleotides

A;Pathway: protein biosynthesis

C;Keywords: blocked amino end; GTP; protein biosynthesis

F;107/Binding site: magnesium (Asp) #status predicted

Query Match 82.9%; Score 29; DB 2; Length 452;  
Best Local Similarity 75.0%; Pred. No. 90;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
||| ||:|  
Db 415 VEAGAE IR 422

RESULT 11

F91063

DNA repair protein RecN [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C;Accession: F91063

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F91063  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-553 <HAY>  
A;Cross-references: UNIPROT:Q8X9B9; GB:BA000007; PIDN:BAB36901.1; PID:g13362949;  
GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs3478  
C;Superfamily: recN protein

Query Match 82.9%; Score 29; DB 2; Length 553;  
Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
| | | | |  
Db 278 VEASDEL R 285

RESULT 12

B85908

protein used in recombination and DNA repair [imported] - Escherichia coli  
(strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C;Accession: B85908

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose,  
D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.;  
Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis,  
N.W.; Lim, A.; Dimalanta, E.; Potamosis, K.; Apodaca, J.; Anantharaman, T.S.;  
Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.  
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: B85908

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-553 <STO>

A;Cross-references: UNIPROT:Q8X9B9; GB:AE005174; NID:g12517040; PIDN:AAG57726.1;  
GSPDB:GN00145; UWGP:Z3909

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: recN

C;Superfamily: recN protein

Query Match 82.9%; Score 29; DB 2; Length 553;  
Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
| | | | |  
Db 278 VEASDEL R 285

RESULT 13

T19756

hypothetical protein C35C5.1 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T19756; T21561  
 R;White, S.  
 submitted to the EMBL Data Library, August 1996  
 A;Reference number: Z19173  
 A;Accession: T19756  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-2962 <WIL>  
 A;Cross-references: UNIPROT:Q93326; EMBL:Z78417; PIDN:CAB01693.1; GSPDB:GN00028; CESP:C35C5.1  
 A;Experimental source: clone C35C5  
 R;Harris, B.  
 submitted to the EMBL Data Library, August 1996  
 A;Reference number: Z19441  
 A;Accession: T21561  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-2962 <WI2>  
 A;Cross-references: EMBL:Z78543; PIDN:CAB01756.1; GSPDB:GN00028; CESP:C35C5.1  
 A;Experimental source: clone F29G6  
 C;Genetics:  
 A;Gene: CESP:C35C5.1  
 A;Map position: X  
 A;Introns: 373/3; 452/2; 525/2; 576/3; 598/3; 782/2; 894/1; 937/3; 1334/3; 1360/3; 1587/2; 1914/1; 1943/1; 2038/1; 2327/3; 2454/3; 2551/2; 2573/3; 2888/2

Query Match 82.9%; Score 29; DB 2; Length 2962;  
 Best Local Similarity 87.5%; Pred. No. 6.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 || |||||  
 Db 2155 VEFSAE LR 2162

RESULT 14  
 S60208  
 fomE protein - *Streptomyces wedmorensis*  
 C;Species: *Streptomyces wedmorensis*  
 C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C;Accession: S60208  
 R;Hidaka, T.; Goda, M.; Kuzuyama, T.; Takei, N.; Hidaka, M.; Seto, H.  
 Mol. Gen. Genet. 249, 274-280, 1995  
 A;Title: Cloning and nucleotide sequence of fosfomycin biosynthetic genes of *Streptomyces wedmorensis*.  
 A;Reference number: S60207; MUID:96091152; PMID:7500951  
 A;Accession: S60208  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-233 <HID>  
 A;Cross-references: UNIPROT:P96073; EMBL:D38561; NID:g3452578; PIDN:BAA32489.1; PID:g1061000  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C;Genetics:  
A;Gene: fomE  
A;Start codon: GTG  
C;Superfamily: Escherichia coli hypothetical 25.2K (lysR-araE intergenic region)

Query Match 80.0%; Score 28; DB 2; Length 233;  
Best Local Similarity 85.7%; Pred. No. 78;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
|||:|  
Db 177 EASAQLR 183

#### RESULT 15

B83487

hypothetical protein PA1263 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C;Accession: B83487

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Coltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: B83487

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-314 <STO>

A;Cross-references: UNIPROT:Q9I481; GB:AE004556; GB:AE004091; NID:g9947194;

PIDN:AAG04652.1; GSPDB:GN00131; PASP:PA1263

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA1263

Query Match 80.0%; Score 28; DB 2; Length 314;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAELR 8  
:|||||  
Db 247 LEASAEFR 254

Search completed: February 10, 2005, 15:59:37  
Job time : 13.3944 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 58.3662 Seconds  
 (without alignments)  
 70.188 Million cell updates/sec

Title: US-10-067-484-10  
 Perfect score: 35  
 Sequence: 1 VEASAE LR 8

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : UniProt\_03:\*  
 1: uniprot\_sprot:\*  
 2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES,

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	32	91.4	391	2	Q70IY5	Q70iy5 streptomyce
2	32	91.4	517	2	Q6QWA9	Q6qwa9 azospirillu
3	32	91.4	555	2	Q7UMA7	Q7uma7 rhodopirell
4	32	91.4	656	2	Q9I4I7	Q9i4i7 pseudomonas
5	31	88.6	398	1	KBL_SALTY	P37419 salmonella
6	31	88.6	398	2	Q8Z2F5	Q8z2f5 salmonella
7	31	88.6	458	2	O54755	O54755 rattus norv
8	31	88.6	573	2	Q8KER5	Q8ker5 chlorobium
9	31	88.6	576	1	MOES_MOUSE	P26041 mus musculu
10	31	88.6	576	1	MOES_RAT	O35763 rattus norv
11	31	88.6	625	2	Q823C5	Q823c5 chlamydophi
12	31	88.6	653	2	Q6G1I6	Q6g1i6 bartonella
13	31	88.6	699	1	SRCH_HUMAN	P23327 homo sapien
14	31	88.6	1765	2	Q9VS30	Q9vs30 drosophila
15	31	88.6	1770	2	Q9VS29	Q9vs29 drosophila
16	30	85.7	132	2	Q6MLD9	Q6mld9 bdellovibri
17	30	85.7	280	2	Q89UC6	Q89uc6 bradyrhizob
18	30	85.7	304	2	Q92X19	Q92x19 rhizobium m
19	30	85.7	304	2	Q8UKN0	Q8ukn0 agrobacteri
20	30	85.7	413	2	Q824A1	Q824a1 chlamydophi
21	30	85.7	441	2	Q8XX87	Q8xx87 ralstonia s

22	30	85.7	460	2	Q62DV0	Q62dv0	burkholderi
23	30	85.7	460	2	Q63JE7	Q63je7	burkholderi
24	30	85.7	532	2	Q9AA46	Q9aa46	caulobacter
25	30	85.7	708	2	Q6QW89	Q6qw89	azospirillu
26	30	85.7	711	2	Q7W206	Q7w206	bordetella
27	30	85.7	713	2	Q7VUK9	Q7vuk9	bordetella
28	30	85.7	713	2	Q7WQY4	Q7wqy4	bordetella
29	29	82.9	89	2	Q8TN37	Q8tn37	methanosarc
30	29	82.9	95	2	Q20843	Q20843	caenorhabdi
31	29	82.9	108	2	Q8NPP8	Q8npp8	corynebacte
32	29	82.9	136	2	Q6LZZ4	Q6lzz4	methanococc
33	29	82.9	147	2	Q9X8M0	Q9x8m0	streptomyce
34	29	82.9	210	2	Q82B37	Q82b37	streptomyce
35	29	82.9	211	2	Q6XLM6	Q6xlm6	brassica ol
36	29	82.9	212	2	Q6XLM9	Q6xlm9	brassica ra
37	29	82.9	213	2	Q6XLM1	Q6xlm1	arabidopsis
38	29	82.9	213	2	Q9SCS7	Q9scs7	arabidopsis
39	29	82.9	215	2	Q6XLM4	Q6xlm4	brassica ol
40	29	82.9	215	2	Q6XLM7	Q6xlm7	brassica ra
41	29	82.9	216	2	Q6XLM5	Q6xlm5	brassica ol
42	29	82.9	216	2	Q6XLM8	Q6xlm8	brassica ra
43	29	82.9	220	2	Q669E8	Q669e8	yersinia ps
44	29	82.9	228	2	Q67SI3	Q67si3	symbiobacte
45	29	82.9	235	2	Q99GR3	Q99gr3	culex nigri

#### ALIGNMENTS

#### RESULT 1

##### Q70IY5

ID Q70IY5 PRELIMINARY; PRT; 391 AA.  
 AC Q70IY5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Putative glycosyltransferase.  
 GN Name=tbmD;  
 OS Streptomyces tenebrarius.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1933;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 17920;  
 RX PubMed=14757238; DOI=10.1016/S0378-1097(03)00881-4;  
 RA Kharel M.K., Basnet D.E., Lee H.C., Liou K., Woo J.S., Kim B.G.,  
 RA Sohng B.G.;  
 RT "Isolation and characterization of tobramycin biosynthetic gene  
 RT cluster from Streptomyces tenebrarius."  
 RL FEMS Microbiol. Lett. 230:185-190(2004).  
 DR EMBL; AJ579650; CAE22469.1; -.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR001296; Glyco\_trans\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 KW Transferase.



SQ SEQUENCE 391 AA; 43428 MW; ECD91C7D2CD46315 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 391;  
Best Local Similarity 87.5%; Pred. No. 1.5e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
|||:||||  
Db 229 VEAAAELR 236

RESULT 2

Q6QWA9

ID Q6QWA9 PRELIMINARY; PRT; 517 AA.  
AC Q6QWA9;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=pRhico055;  
OS Azospirillum brasilense.  
OG Plasmid 90 MDa.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Rhodospirillaceae; Azospirillum.  
OX NCBI\_TaxID=192;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=15033235; DOI=10.1016/S0378-1097(04)00046-1;  
RA Vanbleu E., Marchal K., Lambrecht M., Mathys J., Vanderleyden J.;  
RT "Annotation of the pRhico plasmid of Azospirillum brasilense reveals  
RT its role in determining the outer surface composition."  
RL FEMS Microbiol. Lett. 232:165-172(2004).  
DR EMBL; AY523972; AAS83091.1; -.  
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM\_bind.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 517 AA; 55981 MW; 1F663E452E35E8C4 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 517;  
Best Local Similarity 87.5%; Pred. No. 1.9e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
|||:||||  
Db 314 VEAAAELR 321

RESULT 3

Q7UMA7

ID Q7UMA7 PRELIMINARY; PRT; 555 AA.  
AC Q7UMA7;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Ketoglutarate semialdehyde dehydrogenase (EC 1.2.1.3).  
GN OrderedLocusNames=RB8941;

OS Rhodopirellula baltica.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL; BX294148; CAD76010.1; -.  
 DR HSSP; Q56694; 1EZ0.  
 DR GO; GO:0004029; F:aldehyde dehydrogenase (NAD) activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR CO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002086; Aldehyd\_dehydrog.  
 DR Pfam; PF00171; Aldedh; 1.  
 DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; UNKNOWN\_1.  
 KW Complete proteome; Oxidoreductase.  
 SQ SEQUENCE 555 AA; 57956 MW; 41778B0A11346C8B CRC64;

Query Match 91.4%; Score 32; DB 2; Length 555;  
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 |||:||||  
 Db 81 VEAAAELR 88

#### RESULT 4

Q9I417

ID Q9I417 PRELIMINARY; PRT; 656 AA.  
 AC Q9I417;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Probable protease.  
 GN OrderedLocusNames=PA1327;  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004562; AAG04716.1; -.  
 DR PIR; C83479; C83479.  
 DR GO; GO:0004519; F:endonuclease activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001604; Endonuclease.  
 DR InterPro; IPR008256; Peptidase\_S1B.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF01223; Endonuclease\_NS; 1.  
 DR PRINTS; PR00839; V8PROTEASE.  
 DR SMART; SM00477; NUC; 1.  
 KW Complete proteome; Protease.  
 SQ SEQUENCE 656 AA; 72505 MW; FD536A3BF62811E0 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 656;  
 Best Local Similarity 87.5%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
 |||||:|  
 Db 589 VEASAD LR 596

# RESULT 5

## KBL\_SALTY

ID KBL\_SALTY STANDARD; PRT; 398 AA.  
 AC P37419;  
 DT 01-OCT-1994 (Rel. 30, Created).  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE 2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29) (AKB ligase)  
 DE (Glycine acetyltransferase).  
 GN Name=kbl; OrderedLocusNames=STM3709;  
 OS *Salmonella typhimurium*.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; *Salmonella*.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 RT LT2.";  
 RL Nature 413:852-856(2001).  
 RN [2]  
 RP SEQUENCE OF 1-162 FROM N.A.  
 RC STRAIN=LT2;  
 RX MEDLINE=94209240; PubMed=8157607;

RA Sirisena D.M., Maclachlan P.R., Liu S.L., Hessel A., Sanderson K.E.;  
 RT "Molecular analysis of the rfaD gene, for heptose synthesis, and the  
 RT rfaF gene, for heptose transfer, in lipopolysaccharide synthesis in  
 RT Salmonella typhimurium.";  
 RL J. Bacteriol. 176:2379-2385(1994).  
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + glycine = CoA + 2-amino-3-  
 CC oxobutanoate.  
 CC -!- COFACTOR: Pyridoxal phosphate. .  
 CC -!- SIMILARITY: Belongs to the class-II pyridoxal-phosphate-dependent  
 CC aminotransferase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE008872; AAL22568.1; -.  
 DR EMBL; U06472; AAA59063.1; -.  
 DR HSSP; P07912; 1FC4.  
 DR StyGene; SG10191; kbl.  
 DR InterPro; IPR004839; Aminotrans\_I/II.  
 DR InterPro; IPR001917; Aminotrans\_II.  
 DR Pfam; PF00155; Aminotran\_1\_2; 1.  
 DR TIGRFAMs; TIGR01822; 2am3keto\_CoA; 1.  
 DR TIGRFAMs; TIGR01821; 5aminolev\_synth; 1.  
 DR TIGRFAMs; TIGR01825; gly\_Cac\_T\_rel; 1.  
 DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; 1.  
 KW Acyltransferase; Complete proteome; Pyridoxal phosphate; Transferase.  
 FT BINDING 244 244 Pyridoxal phosphate (By similarity).  
 SQ SEQUENCE 398 AA; 43031 MW; 2A18126C63A7A9A5 CRC64;

Query Match 88.6%; Score 31; DB 1; Length 398;  
 Best Local Similarity 87.5%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
 ||| ||||  
 Db 292 VEAGAE LR 299

# RESULT 6

Q8Z2F5

ID Q8Z2F5 PRELIMINARY; PRT; 398 AA.  
 AC Q8Z2F5; Q7C669;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE 2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29).  
 GN Name=kbl; OrderedLocusNames=STY4086, t3810;  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RT enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ty2 / ATCC 700931;  
 RX MEDLINE=22531367; PubMed=12644504;  
 RX DOI=10.1128/JB.185.7.2330-2337.2003;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;  
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
 RT and CT18.";  
 RL J. Bacteriol. 185:2330-2337(2003).  
 DR EMBL; AL627280; CAD03285.1; -.  
 DR EMBL; AE016847; AAO71292.1; -.  
 DR HSSP; P07912; 1FC4.  
 DR GO; GO:0008415; F:acyltransferase activity; IEA.  
 DR GO; GO:0008890; F:glycine C-acetyltransferase activity; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0008483; F:transaminase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR Pfam; PF00155; Aminotran\_1\_2; 1.  
 DR TIGRFAMs; TIGR01822; 2am3keto\_CoA; 1.  
 DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; UNKNOWN\_1.  
 KW Acyltransferase; Complete proteome; Ligase; Transferase.  
 SQ SEQUENCE 398 AA; 43059 MW; 7E8A5B46D57CC417 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 398;  
 Best Local Similarity 87.5%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VEASAE LR 8  
 ||| |||  
 Db 292 VEAGAE LR 299

#### RESULT 7

O54755

ID O54755 PRELIMINARY; PRT; 458 AA.  
 AC O54755;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE MIPP65.  
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97428277; PubMed=9281354; DOI=10.1006/excr.1997.3658;  
 RA Kitagawa M., Mukai H., Ono Y.;  
 RT "Molecular cloning and characterization of a novel mitochondrial  
 RT phosphoprotein, MIPP65, from rat liver."  
 RL Exp. Cell Res. 235:71-78(1997).  
 DR EMBL; AB000098; BAA24351.1; -.  
 SQ SEQUENCE 458 AA; 49311 MW; C311A2A3F503412F CRC64;

Query Match 88.6%; Score 31; DB 2; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAE LR 8  
 |||||  
 Db 298 EASAE LR 304

# RESULT 8

## Q8KER5

ID Q8KER5 PRELIMINARY; PRT; 573 AA.  
 AC Q8KER5;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Acetolactate synthase, large subunit.  
 GN Name=ilvB; OrderedLocusNames=CT0618;  
 OS Chlorobium tepidum.  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorobaculum.  
 OX NCBI\_TaxID=1097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;  
 RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;  
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,  
 RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,  
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,  
 RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,  
 RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,  
 RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,  
 RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;  
 RT "The complete genome sequence of Chlorobium tepidum TLS, a  
 RT photosynthetic, anaerobic, green-sulfur bacterium."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).  
 DR EMBL; AE012835; AAM71860.1; -.  
 DR HSSP; P07342; 1NOH.  
 DR TIGR; CT0618; -.  
 DR GO; GO:0003984; F:acetolactate synthase activity; IEA.  
 DR GO; GO:0009082; P:branched chain family amino acid biosynthesis; IEA.  
 DR InterPro; IPR004407; Acolac\_synthlrg.  
 DR InterPro; IPR000399; Pyruvate\_decarb.

DR Pfam; PF00205; TPP\_enzyme\_M; 1.  
 DR Pfam; PF02776; TPP\_enzyme\_N; 1.  
 DR TIGRFAMs; TIGR00118; acolac\_lg; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 573 AA; 62874 MW; 164DC5949E86C26E CRC64;

Query Match 88.6%; Score 31; DB 2; Length 573;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAE LR 8  
 |||||  
 Db 223 EASAE LR 229

# RESULT 9

## MOES\_MOUSE

ID MOES\_MOUSE STANDARD; PRT; 576 AA.  
 AC P26041; Q8BSN4;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Moesin (Membrane-organizing extension spike protein).  
 GN Name=Msn;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93055012; PubMed=1429901;  
 RA Sato N., Funayama N., Nagafuchi A., Yonemura S., Tsukita S.,  
 RA Tsukita S.;  
 RT "A gene family consisting of ezrin, radixin and moesin. Its specific  
 RT localization at actin filament/plasma membrane association sites.";  
 RL J. Cell Sci. 103:131-143(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J, and NOD; TISSUE=Forelimb, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."; -  
 RL Nature 420:563-573 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Breast tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF 11-576 FROM N.A.  
 RX MEDLINE=92243764; PubMed=1573844;  
 RA Furthmayr H., Lankes W.T., Amieva M.R.;  
 RT "Moesin, a new cytoskeletal protein and constituent of filopodia: its  
 RT role in cellular functions.";  
 RL Kidney Int. 41:665-670 (1992).  
 RN [5]  
 RP PHOSPHORYLATION SITE SER-575.  
 RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;  
 RA Shu H., Chen S., Bi Q., Mumby M., Brekken D.L.;  
 RT "Identification of phosphoproteins and their phosphorylation sites in  
 RT the WEHI-231 B lymphoma cell line.";  
 RL Mol. Cell. Proteomics 3:279-286 (2004).  
 CC -!- FUNCTION: Probably involved in connections of major cytoskeletal  
 CC structures to the plasma membrane.  
 CC -!- SUBUNIT: Binds SLC9A3R1 (By similarity).  
 CC -!- SIMILARITY: Contains 1 FERM domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; S47577; AAA11762.1; -.  
DR EMBL; AK031171; BAC27288.1; -.  
DR EMBL; AK088336; BAC40290.1; -.  
DR EMBL; BC047366; AAH47366.1; -.  
DR EMBL; M86390; AAA39728.1; -.  
DR HSSP; P26038; 1E5W.  
DR MGD; MGI:97167; Msn.  
DR GO; GO:0016324; C:apical plasma membrane; IDA.  
DR GO; GO:0016323; C:basolateral plasma membrane; IDA.  
DR InterPro; IPR000299; Band\_4.1.  
DR InterPro; IPR011174; ERM.  
DR InterPro; IPR000798; Ez/rad/moesin.  
DR InterPro; IPR009065; FERM.  
DR InterPro; IPR008954; Moesin.  
DR InterPro; IPR011036; PH\_related.  
DR Pfam; PF00373; Band\_41; 1.  
DR Pfam; PF00769; ERM; 1.  
DR PIRSF; PIRSF002305; ERM; 1.  
DR PRINTS; PR00935; BAND41.  
DR SMART; SM00295; B41; 1.  
DR PROSITE; PS00660; FERM\_1; 1.  
DR PROSITE; PS00661; FERM\_2; 1.  
DR PROSITE; PS50057; FERM\_3; 1.  
KW Cytoskeleton; Phosphorylation; Structural protein.  
FT INIT\_MET 0 0 By similarity.  
FT DOMAIN 1 294 FERM.  
FT MOD\_RES 575 575 Phosphoserine.  
FT CONFLICT 330 331 EL -> DV (in Ref. 4).  
FT CONFLICT 370 371 RA -> SP (in Ref. 4).  
SQ SEQUENCE 576 AA; 67635 MW; 5E0F4555552E9145 CRC64;

Query Match 88.6%; Score 31; DB 1; Length 576;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAELR 8  
|||  
Dh 488 EASAELR 494

#### RESULT 10

##### MOES\_RAT

ID MOES\_RAT STANDARD; PRT; 576 AA.  
AC O35763;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Moesin (Membrane-organizing extension spike protein).  
GN Name=Msn;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PHOSPHORYLATION.  
 RX MEDLINE=20404063; PubMed=10945828;  
 RA Theoharides T.C., Wang L., Pang X., Letourneau R., Culm K.E., Basu S.,  
 RA Wang Y., Correia I.;  
 RT "Cloning and cellular localization of the rat mast cell 78-kDa protein  
 RT phosphorylated in response to the mast cell 'stabilizer' cromolyn.";  
 RL J. Pharmacol. Exp. Ther. 294:810-821(2000).  
 CC -!- FUNCTION: Probably involved in connections of major cytoskeletal  
 CC structures to the plasma membrane.  
 CC -!- SUBUNIT: Binds SLC9A3R1 (By similarity).  
 CC -!- PTM: Phosphorylated.  
 CC -!- SIMILARITY: Contains 1 FERM domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF004811; AAB61666.1; -.  
 DR HSSP; P26038; 1E5W.  
 DR RGD; 621260; Msn.  
 DR InterPro; IPR000299; Band\_4.1.  
 DR InterPro; IPR011174; ERM.  
 DR InterPro; IPR000798; Ez/rad/moesin.  
 DR InterPro; IPR009065; FERM.  
 DR InterPro; IPR008954; Moesin.  
 DR InterPro; IPR011036; PH\_related.  
 DR Pfam; PF00373; Band\_41; 1.  
 DR Pfam; PF00769; ERM; 1.  
 DR PIRSF; PIRSF002305; ERM; 1.  
 DR PRINTS; PR00935; BAND41.  
 DR SMART; SM00295; B41; 1.  
 DR PROSITE; PS00660; FERM\_1; 1.  
 DR PROSITE; PS00661; FERM\_2; 1.  
 DR PROSITE; PS50057; FERM\_3; 1.  
 KW Cytoskeleton; Phosphorylation; Structural protein.  
 FT INIT\_MET 0 0 By similarity.  
 FT DOMAIN 1 294 FERM.  
 FT MOD\_RES 575 575 Phosphoserine (By similarity).  
 SQ SEQUENCE 576 AA; 67607 MW; 59606907B2D89938 CRC64;

Query Match 88.6%; Score 31; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
 |||||  
 Db 488 EASAELR 494

RESULT 11

## Q823C5

ID Q823C5 PRELIMINARY; PRT; 625 AA.  
AC Q823C5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocusNames=CCA00500;  
OS Chlamydophila caviae.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GPIC;  
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;  
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,  
RA Heidelberg J.F., Holtzapple E.K., Khouri H.M., Federova N.B.,  
RA Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,  
RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,  
RA Bavoil P.M., Fraser C.M.;  
RT "Genome sequence of Chlamydophila caviae (Chlamydia psittaci GPIC):  
RT examining the role of niche-specific genes in the evolution of the  
RT Chlamydiaceae.";  
RL Nucleic Acids Res. 31:2134-2147(2003).  
DR EMBL; AE016995; AAP05244.1; -.  
DR TIGR; CCA00500; -.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 625 AA; 68389 MW; 0CEC423A69166619 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 625;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
|||||  
Db 101 EASAELR 107

## RESULT 12

## Q6G1I6

ID Q6G1I6 PRELIMINARY; PRT; 653 AA.  
AC Q6G1I6;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocusNames=BQ00400;  
OS Bartonella quintana (Rochalimaea quintana).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bartonellaceae; Bartonella.  
OX NCBI\_TaxID=803;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Toulouse;  
RX PubMed=15210978; DOI=10.1073/pnas.0305659101;  
RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,  
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,

RA La Scola B., Holmberg M., Andersson S.G.E.;  
 RT "The louse-borne human pathogen Bartonella quintana is a genomic  
 RT derivative of the zoonotic agent Bartonella henselae.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).  
 DR EMBL; BX897700; CAF25547.1; -.  
 DR InterPro; IPR000184; Bac\_surfAg\_D15.  
 DR Pfam; PF01103; Bac\_surface\_Ag; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 653 AA; 71355 MW; B2D01CD50B7748F4 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 653;  
 Best Local Similarity 87.5%; Pred. No. 4e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEASAE LR 8  
 || |||||  
 Db 572 VEGSAE LR 579

# RESULT 13

## SRCH\_HUMAN

ID SRCH\_HUMAN STANDARD; PRT; 699 AA.  
 AC P23327;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Sarcoplasmic reticulum histidine-rich calcium-binding protein  
 DE precursor.  
 GN Name=HRC; Synonyms=HCP;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=91244309; PubMed=2037293;  
 RA Hofmann S.L., Topham M., Hsieh C.-L., Francke U.;  
 RT "cDNA and genomic cloning of HRC, a human sarcoplasmic reticulum  
 RT protein, and localization of the gene to human chromosome 19 and mouse  
 RT chromosome 7.";  
 RL Genomics 9:656-669(1991).  
 CC -!- FUNCTION: May play a role in the regulation of calcium  
 CC sequestration or release in the SR of skeletal and cardiac muscle.  
 CC -!- SUBCELLULAR LOCATION: Sarcoplasmic reticulum lumen.  
 CC -!- SIMILARITY: Strong, to rabbit HRC.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M60052; AAA88071.1; -.  
 DR PIR; A54660; A54660.

DR Genew; HGNC:5178; HRC.  
 DR MIM; 142705; -.  
 DR GO; GO:0005509; F:calcium ion binding; TAS.  
 DR GO; GO:0006936; P:muscle contraction; TAS.  
 KW Calcium-binding; Polymorphism; Repeat; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 699 Sarcoplasmic reticulum histidine-rich  
 FT calcium-binding protein.  
 FT DOMAIN 193 204 Glu-rich (acidic).  
 FT DOMAIN 246 261 Asp-rich (acidic).  
 FT DOMAIN 106 365 4 X tandem repeats, acidic.  
 FT REPEAT 180 213 1-1.  
 FT REPEAT 238 270 1-2.  
 FT REPEAT 295 318 1-3.  
 FT REPEAT 343 365 1-4.  
 FT DOMAIN 106 342 6 X approximate tandem repeats.  
 FT REPEAT 106 121 2-1.  
 FT REPEAT 134 154 2-2.  
 FT REPEAT 155 177 2-3.  
 FT REPEAT 214 237 2-4.  
 FT REPEAT 271 294 2-5.  
 FT REPEAT 319 342 2-6.  
 FT DOMAIN 627 673 Metal-binding (Potential).  
 FT VARIANT 96 96 S -> A (in dbSNP:3745297).  
 FT /FTId=VAR\_005623.  
 FT VARIANT 204 204 Missing.  
 FT /FTId=VAR\_011622.  
 SQ SEQUENCE 699 AA; 80244 MW; 9922EEDF012C61DD CRC64;

Query Match 88.6%; Score 31; DB 1; Length 699;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EASAE LR 8  
 |||||  
 Db 52 EASAE LR 58

#### RESULT 14

Q9VS30

ID Q9VS30 PRELIMINARY; PRT; 1765 AA.  
 AC Q9VS30;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG32387-PB.  
 GN ORFNames=CG32387;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RT *melanogaster* euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AE003560; AAF50600.2; -.  
 DR HSSP; P13590; 1IE5.  
 DR FlyBase; FBgn0052387; CG32387.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR008957; FN\_III-like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR Pfam; PF00041; fn3; 6.  
 DR Pfam; PF00047; ig; 6.  
 DR SMART; SM00060; FN3; 6.  
 DR SMART; SM00408; IGc2; 8.  
 DR PROSITE; PS50853; FN3; 6.  
 DR PROSITE; PS50835; IG\_LIKE; 9.  
 SQ SEQUENCE 1765 AA; 193273 MW; 81A6A498A903E905 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 1765;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
 |||||  
 Db 284 EASAELR 290

# RESULT 15

Q9VS29

ID Q9VS29 PRELIMINARY; PRT; 1770 AA.  
 AC Q9VS29;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG32387-PA.  
 GN ORFNames=CG32387;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RT *melanogaster* euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.



RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirska R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AE003560; AAF50601.2; -.  
 DR HSSP; P13590; 1IE5.  
 DR FlyBase; FBgn0052387; CG32387.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR008957; FN\_III-like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR Pfam; PF00041; fn3; 6.  
 DR Pfam; PF00047; ig; 6.  
 DR SMART; SM00060; FN3; 6.  
 DR SMART; SM00408; IGc2; 8.  
 DR PROSITE; PS50853; FN3; 6.  
 DR PROSITE; PS50835; IG\_LIKE; 9.  
 SQ SEQUENCE 1770 AA; 193904 MW; 6C1D43D05CC98A7A CRC64;

Query Match 88.6%; Score 31; DB 2; Length 1770;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EASAELR 8  
 |||||  
 Db 284 EASAELR 290

Search completed: February 10, 2005, 15:57:39  
 Job time : 59.3662 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 78.4648 Seconds  
(without alignments)  
44.362 Million cell updates/sec

Title: US-10-067-484-11  
Perfect score: 41  
Sequence: 1 LLSGLSDTV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	41	100.0	9	5	ABB81978	Abb81978 30 kDa ra
2	35	85.4	122	4	AAG89866	Aag89866 C glutami
3	35	85.4	397	5	ABB49314	Abb49314 Listeria
4	35	85.4	553	2	AAW10559	Aaw10559 Barley gi
5	35	85.4	553	8	ADJ49115	Adj49115 Oil-assoc
6	35	85.4	553	8	ADJ50351	Adj50351 Oil-assoc
7	34	82.9	143	3	AAB41908	Aab41908 Human ORF
8	34	82.9	543	4	AAM39906	Aam39906 Human pol
9	34	82.9	565	4	AAM41692	Aam41692 Human pol

10	34	82.9	682	7	ADG39871	Adg39871	Protein s
11	34	82.9	682	7	ADG39872	Adg39872	Protein s
12	34	82.9	708	3	AAy70473	Aay70473	Human cyc
13	34	82.9	708	4	ABU52881	Abu52881	Human ute
14	34	82.9	798	8	ADN24197	Adn24197	Bacterial
15	34	82.9	803	7	ADC15007	Adc15007	Mouse adi
16	34	82.9	803	7	ADC15015	Adc15015	Human adi
17	34	82.9	803	8	ADR09630	Adr09630	Human pro
18	34	82.9	2636	7	ABO77007	Abo77007	Pseudomon
19	34	82.9	5149	6	ABU15709	Abu15709	Protein e
20	33	80.5	61	4	AAU42021	Aau42021	Propionib
21	33	80.5	61	6	ABM38540	Abm38540	Propionib
22	33	80.5	76	6	ABU49462	Abu49462	Protein e
23	33	80.5	113	8	ADS29784	Ads29784	Bacterial
24	33	80.5	167	3	AAG23305	Aag23305	Arabidops
25	33	80.5	221	3	AAG23304	Aag23304	Arabidops
26	33	80.5	255	3	AAG46382	Aag46382	Arabidops
27	33	80.5	260	3	AAG46381	Aag46381	Arabidops
28	33	80.5	282	3	AAG46380	Aag46380	Arabidops
29	33	80.5	349	1	AAP70463	Aap70463	Sequence
30	33	80.5	406	6	ADA35277	Ada35277	Acinetoba
31	33	80.5	518	7	ABO74745	Abo74745	Pseudomon
32	33	80.5	527	8	ADQ66940	Adq66940	Novel hum
33	33	80.5	717	6	ABR59723	Abr59723	Human RIK
34	33	80.5	843	6	ABR59722	Abr59722	Human KIA
35	33	80.5	845	5	ABG92078	Abg92078	Human rec
36	33	80.5	845	8	ADS88196	Ads88196	Human pro
37	33	80.5	1311	2	AAW52197	Aaw52197	Precis co
38	33	80.5	1311	2	AAW72971	Aaw72971	Precis co
39	33	80.5	1311	4	AAB67156	Aab67156	Butterfly
40	33	80.5	1311	5	AAG79573	Aag79573	Butterfly
41	33	80.5	1311	7	ABU62149	Abu62149	Buckeye p
42	33	80.5	1311	7	ADE94209	Ade94209	Peacock b
43	33	80.5	1311	7	ADH62716	Adh62716	Butterfly
44	33	80.5	1311	8	ADE48974	Ade48974	Butterfly
45	33	80.5	1320	4	ABB64922	Abb64922	Drosophil

#### ALIGNMENTS

##### RESULT 1

ABB81978

ID ABB81978 standard; peptide; 9 AA.

XX

AC ABB81978;

XX

DT 25-NOV-2002 (first entry)

XX

DE 30 kDa ragweed pollen allergen tryptic peptide 11.

XX

KW Ragweed; pollen; allergen; Ambt 7; glycoprotein; antiallergic;  
KW immunotherapy; disulphide protein.

XX

OS Ambrosia elatior.

XX

PN WO200263012-A2.

XX  
 PD 15-AUG-2002.  
 XX  
 PF 04-FEB-2002; 2002WO-US003346.  
 XX  
 PR 05-FEB-2001; 2001US-0266686P.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Buchanan BB, Del Val G, Frick OL;  
 XX  
 DR WPI; 2002-657539/70.  
 XX  
 PT New ragweed pollen allergens, useful in allergy testing and immunotherapy  
 PT regimens, particularly for treating sensitivity to pollen or pollen  
 PT allergy (e.g. anaphylaxis, or symptoms of hives or asthma) in a mammal,  
 PT especially a human.  
 XX  
 PS Claim 1; Page 53; 70pp; English.  
 XX  
 CC The invention relates to an isolated pollen allergen purified from  
 CC ragweed pollen, substantially free of any other pollen proteins, or a  
 CC protein that is an antigenic fragment of a pollen allergen Ambt 7. The  
 CC allergen is characterized by the following physiochemical and biological  
 CC properties: (a) being contained in pollen extracts; (b) a glycoprotein;  
 CC (c) a sulphhydryl group containing protein; (d) a molecular weight of  
 CC about 30 kDa as determined by SDS-polyacrylamide gel electrophoresis; and  
 CC (e) possessing allergen activity. The pollen allergen, or antigenic  
 CC protein fragment of the pollen allergen Ambt 7, or composition is useful  
 CC for treating sensitivity to pollen or pollen allergy in a mammal. This  
 CC allergy includes anaphylaxis or atopy, which includes the symptoms of hay  
 CC fever, asthma or hives. The allergen is also useful in allergy testing  
 CC and immunotherapy regimens. Sequences ABB81968-978 represent tryptic  
 CC peptide fragments of the 30 kDa ragweed complete pollen extract  
 CC disulphide protein allergen  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 41; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 |||||  
 Db 1 LLSGLSDTV 9

RESULT 2  
 AAG89866  
 ID AAG89866 standard; protein; 122 AA.  
 XX  
 AC AAG89866;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE C glutamicum protein fragment SEQ ID NO: 3620.  
 XX

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
 KW organic acid synthesis.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN EP1108790-A2.  
 XX  
 PD 20-JUN-2001.  
 XX  
 PF 18-DEC-2000; 2000EP-00127688.  
 XX  
 PR 16-DEC-1999; 99JP-00377484.  
 PR 07-APR-2000; 2000JP-00159162.  
 PR 03-AUG-2000; 2000JP-00280988.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX  
 DR WPI; 2001-376931/40.  
 DR N-PSDB; AAH65085.  
 XX  
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing  
 PT expression profile or pattern of a gene and identifying homologous gene,  
 XX  
 PS Claim 17; SEQ ID NO 3620; 246pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of nucleotide and protein  
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of coryneform bacterium, measuring expression amount and analysing  
 CC the expression profile or expression pattern of a gene derived from  
 CC Coryneform bacterium, and identifying a homologue of a gene derived from  
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described in the  
 CC exemplification of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from the European Patent Office  
 XX  
 SQ Sequence 122 AA;

Query Match 85.4%; Score 35; DB 4; Length 122;  
 Best Local Similarity 87.5%; Pred. No. 47;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
 :|||||  
 Db 52 ISGLSDTV 59

RESULT 3  
 ABB49314  
 ID ABB49314 standard; protein; 397 AA.  
 XX

AC ABB49314;  
 XX  
 DT 05-FEB-2002 (first entry)  
 XX  
 DE Listeria monocytogenes protein #2018.  
 XX  
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KW vitamin B12; bacterial infection; disease.  
 XX  
 OS Listeria monocytogenes.  
 XX  
 PN WO200177335-A2.  
 XX  
 PD 16-OCT-2001.  
 XX  
 PF 11-APR-2001; 2001WO-FR001118.  
 XX  
 PR 11-APR-2000; 2000FR-00004629.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;  
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
 PI Rose M, Voss H;  
 XX  
 DR WPI; 2002-010914/01.  
 XX  
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
 PT and prevention of Listeria and related bacterial infections, and related  
 PT polypeptides.  
 XX  
 PS Claim 6; SEQ ID NO 2019; 192pp; French.  
 XX  
 CC The present invention relates to the genome sequence of Listeria  
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of  
 CC it are useful for selecting probes and primers for detecting genes in L.  
 CC monocytogenes and related organisms, and for studying genetic  
 CC polymorphisms and other genomes. The present sequence is a protein  
 CC encoded by the genome sequence of the present invention. Proteins  
 CC expressed from the genome sequence are useful for raising specific  
 CC antibodies, identification of L. monocytogenes and related organisms, and  
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
 CC B12. The genome sequence and proteins encoded by it are also useful for  
 CC selecting compounds that regulate gene expression and cell replication  
 CC and modulate L. monocytogenes-related diseases. In addition, the genome  
 CC sequence and proteins encoded by it are useful in pharmaceutical and  
 CC vaccines compositions for the treatment or prevention of infections by L.  
 CC monocytogenes and related organisms. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 397 AA;

Query Match 85.4%; Score 35; DB 5; Length 397;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
|| |||||:  
Db 89 LLCGLSDTI 97

RESULT 4

AAW10559

ID AAW10559 standard; protein; 553 AA.

XX

AC AAW10559;

XX

DT 17-OCT-2003 (revised)

DT 21-APR-1997 (first entry)

XX

DE Barley gibberellin-regulated MYB polypeptide.

XX

KW Gibberellin-regulated MYB polypeptide; GAMYB; barley; rice; maize;

KW sorghum; rye; wheat; malting; alpha-amylase; beta-glucanase;

KW alpha-glucosidase; xylanase; cathepsin beta-like protease;

KW arabinofuranosidase; stem elongation; flowering; germination; dormancy;

KW transgenic plant.

XX

OS Hordeum vulgare; strain Himalaya.

XX

FH Key Location/Qualifiers

FT Region 42. .145

FT /label= R2/R3

FT /note= "R2 and R3 motifs are conserved in MYB-related proteins"

XX

PN WO9700961-A1.

XX

PD 09-JAN-1997.

XX

PF 21-JUN-1996; 96WO-AU000383.

XX

PR 23-JUN-1995; 95AU-00003779.

PR 09-NOV-1995; 95AU-00006470.

XX

PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.

PA (AUSU ) UNIV AUSTRALIAN NAT.

XX

PI Gubler FJ, Jacobsen JV;

XX

DR WPI; 1997-087388/08.

DR N-PSDB; AAT60744.

XX

PT Nucleic acid encoding giberellin-regulated polypeptide(s) - which  
PT regulate synthesis of malting enzymes in plants, e.g. barley and rice,  
PT and modify stem elongation, flowering, germination, etc.

XX

PS Claim 21; Page 62-64; 108pp; English.

XX

CC Barley gibberellin-regulated MYB polypeptide GAMYB (AAW10559) regulates  
CC the expression of gibberellin (GA)-regulated genes that encode hydrolytic  
CC malting enzymes such as alpha-amylase, beta-glucanases, cathepsin-like  
CC protease, alpha-glucosidase, xylanase, arabinofuranosidase, etc. It is  
CC encoded by a cDNA clone (AAT60744) obtd. from a barley aleurone cDNA  
CC library. By altering the expression of GAMYB (using e.g. antisense or  
CC ribozyme constructs) in transgenic plants, the expression of GA-regulated  
CC genes involved in stem elongation, flowering, leaf development, fruit set  
CC and growth, germination, sex determination and esp. malting can be  
CC modified. GAMYB peptides (see also AAW10557-58) can be used to raise  
CC antibodies useful for screening for related sequences. (Updated on 17-OCT  
CC -2003 to standardise OS field)

XX

SQ Sequence 553 AA;

Query Match 85.4%; Score 35; DB 2; Length 553;  
Best Local Similarity 77.8%; Pred. No. 2.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
|| |||||:  
Db 247 LLPGLSDTI 255

#### RESULT 5

ADJ49115

ID ADJ49115 standard; protein; 553 AA.

XX

AC ADJ49115;

XX

DT 06-MAY-2004 (first entry)

XX

DE Oil-associated gene related protein #615.

XX

KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.

XX

OS Unidentified.

XX

PN US2004025202-A1.

XX

PD 05-FEB-2004.

XX

PF 14-MAR-2003; 2003US-00389566.

XX

PR 15-MAR-2002; 2002US-0365301P.

PR 26-JUN-2002; 2002US-0391786P.

PR 26-JUN-2002; 2002US-0392018P.

XX

PA (LAUR/) LAURIE C C.

PA (RAVA/) RAVANELLO M.

PA (SAVA/) SAVAGE T.

PA (LEDE/) LEDEAUX J R.

PA (ROGE/) ROGERS J A.

XX

PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;



XX  
 DR WPI; 2004-142683/14.  
 XX  
 PT Novel recombinant DNA construct comprising a promoter functional in  
 PT plants operably linked to an oil-associated gene for producing transgenic  
 PT plant seed.  
 XX  
 PS Example 3; SEQ ID NO 1119; 22pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in plants operably linked to an oil-associated gene.  
 CC The construct is useful for transgenic plant seed which has in its genome  
 CC the construct, that is functional in the plant to transcribe the oil-  
 CC associated gene. The transgenic plant seed grows into a plant having  
 CC enhanced seed oil as compared to wild type. The construct is useful for  
 CC producing hybrid maize seed. The transgenic plant seed is useful for  
 CC producing vegetable oil. The present sequence represents the amino acid  
 CC sequence of an oil-associated gene related protein.  
 XX  
 SQ Sequence 553 AA;

Query Match 85.4%; Score 35; DB 8; Length 553;  
 Best Local Similarity 77.8%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 || |||||:  
 Db 247 LLPGLSDTI 255

# RESULT 6

ADJ50351

ID ADJ50351 standard; protein; 553 AA.

XX

AC ADJ50351;

XX

DT 06-MAY-2004 (first entry)

XX

DE Oil-associated gene related protein #1851.

XX

KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.

XX

OS Unidentified.

XX

PN US2004025202-A1.

XX

PD 05-FEB-2004.

XX

PF 14-MAR-2003; 2003US-00389566.

XX

PR 15-MAR-2002; 2002US-0365301P.

PR 26-JUN-2002; 2002US-0391786P.

PR 26-JUN-2002; 2002US-0392018P.

XX

PA (LAUR/) LAURIE C C.

PA (RAVA/) RAVANELLO M.

PA (SAVA/) SAVAGE T.

PA (LEDE/) LEDEAUX J R.  
 PA (ROGE/) ROGERS J A.  
 XX  
 PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;  
 XX  
 DR WPI; 2004-142683/14.  
 XX  
 PT Novel recombinant DNA construct comprising a promoter functional in  
 PT plants operably linked to an oil-associated gene for producing transgenic  
 PT plant seed.  
 XX  
 PS Example 3; SEQ ID NO 2355; 22pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in plants operably linked to an oil-associated gene.  
 CC The construct is useful for transgenic plant seed which has in its genome  
 CC the construct, that is functional in the plant to transcribe the oil-  
 CC associated gene. The transgenic plant seed grows into a plant having  
 CC enhanced seed oil as compared to wild type. The construct is useful for  
 CC producing hybrid maize seed. The transgenic plant seed is useful for  
 CC producing vegetable oil. The present sequence represents the amino acid  
 CC sequence of an oil-associated gene related protein.  
 XX  
 SQ Sequence 553 AA;

Query Match 85.4%; Score 35; DB 8; Length 553;  
 Best Local Similarity 77.8%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
 || |||||:  
 Db 247 LLPGLSDTI 255

# RESULT 7

AAB41908

ID AAB41908 standard; protein; 143 AA.

XX

AC AAB41908;

XX

DT 08-FEB-2001 (first entry)

XX

DE Human ORFX ORF1672 polypeptide sequence SEQ ID NO:3344.

XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;

KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach M;  
 XX  
 DR WPI; 2000-602362/57.  
 DR N-PSDB; AAC76117.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 2532; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiatic; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 143 AA;

Query Match 82.9%; Score 34; DB 3; Length 143;  
 Best Local Similarity 87.5%; Pred. No. 90;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDT 8  
 ||||:|||

## RESULT 8

AAM39906

ID AAM39906 standard; protein; 543 AA.

XX

AC AAM39906;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 3051.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 25-DEC-2000; 2000WO-US034263.

XX

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR N-PSDB; AAI59062.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.

XX

PS Example 4; SEQ ID NO 3051; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 543 AA;

Query Match 82.9%; Score 34; DB 4; Length 543;  
Best Local Similarity 77.8%; Pred. No. 3.9e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
:|||| |||  
Db 172 MLSSGLPDTV 180

RESULT 9

AAM41692

ID AAM41692 standard; protein; 565 AA.

XX

AC AAM41692;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 6623.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-US034263.

XX

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX

PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI60848.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Example 2; SEQ ID NO 6623; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX  
 SQ Sequence 565 AA;

Query Match 82.9%; Score 34; DB 4; Length 565;  
 Best Local Similarity 77.8%; Pred. No. 4.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 :|||| |||  
 Db 194 MLSGLPDTV 202

# RESULT 10

ADG39871

ID ADG39871 standard; protein; 682 AA.

XX

AC ADG39871;

XX

DT 26-FEB-2004 (first entry)

XX

DE Protein similar to human NOV15 #2.

XX

KW Human; NOVX; cardiomyopathy; atherosclerosis; hypertension;  
 KW congenital heart defect; pulmonary stenosis; scleroderma; obesity;  
 KW metabolic disturbance; obesity; transplantation; adrenoleukodystrophy;  
 KW congenital adrenal hyperplasia; prostate cancer; diabetes;  
 KW metabolic disorder; neoplasm; adenocarcinoma; fertility; haemophilia;  
 KW graft versus host disease; AIDS; bronchial asthma; Crohn's disease;

KW multiple sclerosis; infectious disease; anorexia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003203843-A1.  
 XX  
 PD 30-OCT-2003.  
 XX  
 PF 11-APR-2002; 2002US-00120801.  
 XX  
 PR 20-APR-2001; 2001US-0285609P.  
 PR 23-APR-2001; 2001US-0285748P.  
 PR 24-APR-2001; 2001US-0286068P.  
 PR 25-APR-2001; 2001US-0286292P.  
 PR 03-MAY-2001; 2001US-0288334P.  
 PR 16-MAY-2001; 2001US-0291241P.  
 PR 14-SEP-2001; 2001US-0322284P.  
 XX  
 PA (PENA/) PENA C E A.  
 PA (GUOX/) GUO X.  
 PA (SHIM/) SHIMKETS R A.  
 PA (PADI/) PADIGARU M.  
 PA (KEKU/) KEKUDA R.  
 PA (SPYT/) SPYTEK K A.  
 PA (MEHR/) MEHRABAN F.  
 PA (TOPP/) TOPPER J N.  
 PA (MALY/) MALYANKAR U M.  
 PA (WASS/) WASSERMAN S M.  
 PA (EDIN/) EDINGER S R.  
 PA (SMIT/) SMITHSON G.  
 PA (GUNT/) GUNTHER E.  
 PA (KOMU/) KOMUVES L.  
 XX  
 PI Pena CEA, Guo X, Shimkets RA, Padigaru M, Kekuda R, Spytek KA;  
 PI Mehraban F, Topper JN, Malyankar UM, Wasserman SM, Edinger SR;  
 PI Smithson G, Gunther E, Komuves L;  
 XX  
 DR WPI; 2003-900671/82.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing or  
 PT treating e.g. cardiomyopathy, atherosclerosis, hypertension, scleroderma,  
 PT obesity, prostate cancer, AIDS, bronchial asthma, Crohn's disease, or  
 PT multiple sclerosis.  
 XX  
 PS Disclosure; SEQ ID NO 103; 215pp; English.  
 XX  
 CC The invention relates to a new isolated polypeptide comprising an amino  
 CC acid sequence selected from 17 fully defined human NOVX sequences (even  
 CC SEQ ID NOS between ADG39770 and ADG39802), a mature form of the NOVX  
 CC amino acid or a variant of NOVX, where one or more amino acid residue in  
 CC the variant differs in no more than 15% of the amino acid residues of  
 CC NOVX. Also included are an isolated nucleic acid (NA) molecule  
 CC (comprising a nucleic acid sequence encoding a NOVX polypeptide above  
 CC (odd SEQ ID NOS between ADG39769 and ADG39801), a nucleic acid fragment

CC encoding at least a portion of a NOVX polypeptide and a complement of  
 CC NOVX NA), a vector comprising NOVX NA, a cell comprising the vector, an  
 CC antibody that immunospecifically binds to NOVX, a method for determining  
 CC the presence or amount of NOVX or NOVX NA in a sample, a method of  
 CC identifying an agent that binds to NOVX, a method for identifying an  
 CC agent that modulates the expression or activity of NOVX, a method for  
 CC modulating the activity of NOVX, a method of treating or preventing a  
 CC NOVX-associated disorder, a method for screening for a modulator of  
 CC activity or of latency or predisposition to a NOVX-associated disorder, a  
 CC method for determining the presence of or predisposition to a disease  
 CC associated with altered levels of NOVX or NOVX NA in a first mammalian  
 CC subject and a method of treating a pathological state in a mammal by  
 CC administering NOVX or an antibody that binds to NOVX. The NOVX  
 CC polypeptide, nucleic acid or antibody is useful in the manufacture of a  
 CC medicament for treating a syndrome associated with a human disease or a  
 CC NOVX-associated disorder. The NOVX polypeptides and nucleic acids  
 CC encoding them are useful for diagnosing or treating pathologies, diseases  
 CC or conditions associated with NOVX sequences, including cardiomyopathy,  
 CC atherosclerosis, hypertension, congenital heart defects, pulmonary  
 CC stenosis, scleroderma, obesity, metabolic disturbances associated with  
 CC obesity, transplantation, adrenoleukodystrophy, congenital adrenal  
 CC hyperplasia, prostate cancer, diabetes, metabolic disorders, neoplasm,  
 CC adenocarcinoma, fertility, haemophilia, graft versus host disease, AIDS,  
 CC bronchial asthma, Crohn's disease, multiple sclerosis, infectious  
 CC disease, anorexia, neurodegenerative disorders (e.g. Alzheimer's disease,  
 CC or Parkinson's disease), immune disorders, haematopoietic disorders,  
 CC dyslipidaemias, and wasting disorders associated with chronic diseases.  
 CC The polypeptides can be used as immunogens to produce antibodies and as  
 CC vaccines. The sequences may further be used in chromosome mapping,  
 CC identifying individual from minute biological samples (tissue typing),  
 CC and in forensic identification of a biological sample. The present  
 CC sequence is a protein showing sequence similarity to a NOVX protein.  
 XX  
 SQ Sequence 682 AA;

Query Match 82.9%; Score 34; DB 7; Length 682;  
 Best Local Similarity 77.8%; Pred. No. 5.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 : |||| |||  
 Db 311 MSLGLPDTV 319

# RESULT 11

ADG39872

ID ADG39872 standard; protein; 682 AA.

XX

AC ADG39872;

XX

DT 26-FEB-2004 (first entry)

XX

DE Protein similar to human NOV15 #3.

XX

KW Mouse; NOVX; cardiomyopathy; atherosclerosis; hypertension;

KW congenital heart defect; pulmonary stenosis; scleroderma; obesity;

KW metabolic disturbance; obesity; transplantation; adrenoleukodystrophy;



KW congenital adrenal hyperplasia; prostate cancer; diabetes;  
 KW metabolic disorder; neoplasm; adenocarcinoma; fertility; haemophilia;  
 KW graft versus host disease; AIDS; bronchial asthma; Crohn's disease;  
 KW multiple sclerosis; infectious disease; anorexia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia;  
 KW wasting disorder.  
 XX  
 OS Mus musculus.  
 XX  
 PN US2003203843-A1.  
 XX  
 PD 30-OCT-2003.  
 XX  
 PF 11-APR-2002; 2002US-00120801.  
 XX  
 PR 20-APR-2001; 2001US-0285609P.  
 PR 23-APR-2001; 2001US-0285748P.  
 PR 24-APR-2001; 2001US-0286068P.  
 PR 25-APR-2001; 2001US-0286292P.  
 PR 03-MAY-2001; 2001US-0288334P.  
 PR 16-MAY-2001; 2001US-0291241P.  
 PR 14-SEP-2001; 2001US-0322284P.  
 XX  
 PA (PENA/) PENA C E A.  
 PA (GUOX/) GUO X.  
 PA (SHIM/) SHIMKETS R A.  
 PA (PADI/) PADIGARU M.  
 PA (KEKU/) KEKUDA R.  
 PA (SPYT/) SPYTEK K A.  
 PA (MEHR/) MEHRABAN F.  
 PA (TOPP/) TOPPER J N.  
 PA (MALY/) MALYANKAR U M.  
 PA (WASS/) WASSERMAN S M.  
 PA (EDIN/) EDINGER S R.  
 PA (SMIT/) SMITHSON G.  
 PA (GUNT/) GUNTHER E.  
 PA (KOMU/) KOMUVES L.  
 XX  
 PI Pena CEA, Guo X, Shimkets RA, Padigaru M, Kekuda R, Spytek KA;  
 PI Mehraban F, Topper JN, Malyankar UM, Wasserman SM, Edinger SR;  
 PI Smithson G, Gunther E, Komuves L;  
 XX  
 DR WPI; 2003-900671/82.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing or  
 PT treating e.g. cardiomyopathy, atherosclerosis, hypertension, scleroderma,  
 PT obesity, prostate cancer, AIDS, bronchial asthma, Crohn's disease, or  
 PT multiple sclerosis.  
 XX  
 PS Disclosure; SEQ ID NO 104; 215pp; English.  
 XX  
 CC The invention relates to a new isolated polypeptide comprising an amino  
 CC acid sequence selected from 17 fully defined human NOVX sequences (even  
 CC SEQ ID NOS between ADG39770 and ADG39802), a mature form of the NOVX  
 CC amino acid or a variant of NOVX, where one or more amino acid residue in  
 CC the variant differs in no more than 15% of the amino acid residues of

CC NOVX. Also included are an isolated nucleic acid (NA) molecule  
 CC (comprising a nucleic acid sequence encoding a NOVX polypeptide above  
 CC (odd SEQ ID NOS between ADG39769 and ADG39801), a nucleic acid fragment  
 CC encoding at least a portion of a NOVX polypeptide and a complement of  
 CC NOVX NA), a vector comprising NOVX NA, a cell comprising the vector, an  
 CC antibody that immunospecifically binds to NOVX, a method for determining  
 CC the presence or amount of NOVX or NOVX NA in a sample, a method of  
 CC identifying an agent that binds to NOVX, a method for identifying an  
 CC agent that modulates the expression or activity of NOVX, a method for  
 CC modulating the activity of NOVX, a method of treating or preventing a  
 CC NOVX-associated disorder, a method for screening for a modulator of  
 CC activity or of latency or predisposition to a NOVX-associated disorder, a  
 CC method for determining the presence of or predisposition to a disease  
 CC associated with altered levels of NOVX or NOVX NA in a first mammalian  
 CC subject and a method of treating a pathological state in a mammal by  
 CC administering NOVX or an antibody that binds to NOVX. The NOVX  
 CC polypeptide, nucleic acid or antibody is useful in the manufacture of a  
 CC medicament for treating a syndrome associated with a human disease or a  
 CC NOVX-associated disorder. The NOVX polypeptides and nucleic acids  
 CC encoding them are useful for diagnosing or treating pathologies, diseases  
 CC or conditions associated with NOVX sequences, including cardiomyopathy,  
 CC atherosclerosis, hypertension, congenital heart defects, pulmonary  
 CC stenosis, scleroderma, obesity, metabolic disturbances associated with  
 CC obesity, transplantation, adrenoleukodystrophy, congenital adrenal  
 CC hyperplasia, prostate cancer, diabetes, metabolic disorders, neoplasm,  
 CC adenocarcinoma, fertility, haemophilia, graft versus host disease, AIDS,  
 CC bronchial asthma, Crohn's disease, multiple sclerosis, infectious  
 CC disease, anorexia, neurodegenerative disorders (e.g. Alzheimer's disease,  
 CC or Parkinson's disease), immune disorders, haematopoietic disorders,  
 CC dyslipidaemias, and wasting disorders associated with chronic diseases.  
 CC The polypeptides can be used as immunogens to produce antibodies and as  
 CC vaccines. The sequences may further be used in chromosome mapping,  
 CC identifying individual from minute biological samples (tissue typing),  
 CC and in forensic identification of a biological sample. The present  
 CC sequence is a protein showing sequence similarity to a NOVX protein.  
 XX  
 SQ Sequence 682 AA;

Query Match 82.9%; Score 34; DB 7; Length 682;  
 Best Local Similarity 77.8%; Pred. No. 5.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 :||| |||  
 Db 311 MSLGLPDTV 319

RESULT 12  
 AAY70473  
 ID AAY70473 standard; protein; 708 AA.  
 XX  
 AC AAY70473;  
 XX  
 DT 04-JUL-2000 (first entry)  
 XX  
 DE Human cyclic nucleotide-associated protein-1 (CNAP-1).  
 XX

KW Cyclic nucleotide-associated protein-1; CNAP-1; human; cytostatic;  
 KW anti-arteriosclerotic; hepatotropic; anti-leukaemic; anti-inflammatory;  
 KW immunomodulatory; anti-asthmatic; anti-anaemic; anti-diabetic; diagnosis;  
 KW anti-sclerotic; dermatological; neuroprotective; anti-epileptic; cancer;  
 KW anti-Alzheimer's; anti-Parkinsonian; cerebroprotective; ophthalmological;  
 KW anti-infertility; anti-allergic; vasotropic; immunosuppressive;  
 KW hypotensive; gene therapy; prevention; treatment; arteriosclerosis;  
 KW cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;  
 KW neurological; vision; reproductive; smooth muscle.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	Peptide	1. .46
FT		/label= Signal_peptide
FT	Modified-site	6
FT		/note= "Potential phosphorylation site"
FT	Modified-site	13
FT		/note= "Potential phosphorylation site"
FT	Protein	47. .708
FT		/label= Mature_CNAP-1
FT		/note= "Shares 15% identity to Saccharomyces kluyveri adenylyl cyclase"
FT	Modified-site	57
FT		/note= "Potential phosphorylation site"
FT	Modified-site	58
FT		/note= "Potential phosphorylation site"
FT	Modified-site	76
FT		/note= "Potential phosphorylation site"
FT	Modified-site	84
FT		/note= "Potential phosphorylation site"
FT	Modified-site	99
FT		/note= "N-glycosylated"
FT	Modified-site	106
FT		/note= "Potential phosphorylation site"
FT	Modified-site	117
FT		/note= "Potential phosphorylation site"
FT	Modified-site	120
FT		/note= "Potential phosphorylation site"
FT	Modified-site	210
FT		/note= "N-glycosylated"
FT	Modified-site	214
FT		/note= "N-glycosylated"
FT	Domain	229. .248
FT		/label= Transmembrane_domain
FT	Modified-site	250
FT		/note= "Potential phosphorylation site"
FT	Modified-site	263
FT		/note= "Potential phosphorylation site"
FT	Modified-site	290
FT		/note= "Potential phosphorylation site"
FT	Modified-site	344
FT		/note= "Potential phosphorylation site"
FT	Modified-site	353
FT		/note= "Potential phosphorylation site"
FT	Region	371. .384
FT		/label= Leucine-rich_repeat

FT		/note= "Signature sequence"
FT	Modified-site	383
FT		/note= "Potential phosphorylation site"
FT	Modified-site	402
FT		/note= "Potential phosphorylation site"
FT	Region	418. .431
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	425
FT		/note= "Potential phosphorylation site"
FT	Modified-site	437
FT		/note= "N-glycosylated"
FT	Modified-site	442
FT		/note= "Potential phosphorylation site"
FT	Region	446. .459
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	448
FT		/note= "Potential phosphorylation site"
FT	Modified-site	453
FT		/note= "Potential phosphorylation site"
FT	Modified-site	457
FT		/note= "N-glycosylated"
FT	Region	493. .506
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	494
FT		/note= "Potential phosphorylation site"
FT	Modified-site	495
FT		/note= "N-glycosylated"
FT	Region	496. .509
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	497
FT		/note= "Potential phosphorylation site"
FT	Region	516. .529
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	518
FT		/note= "Potential phosphorylation site"
FT	Region	519. .532
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	531
FT		/note= "Potential phosphorylation site"
FT	Region	541. .554
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Region	544. .557
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	545
FT		/note= "Potential phosphorylation site"
FT	Region	564. .577
FT		/label= Leucine-rich_repeat
FT		/note= "Signature sequence"
FT	Modified-site	565

FT /note= "Potential phosphorylation site"  
 FT Region 567. .580  
 FT /label= Leucine-rich\_repeat  
 FT /note= "Signature sequence"  
 FT Region 587. .600  
 FT /label= Leucine-rich\_repeat  
 FT /note= "Signature sequence"  
 FT Region 590. .603  
 FT /label= Leucine-rich\_repeat  
 FT /note= "Signature sequence"  
 FT Modified-site 596  
 FT /note= "Potential phosphorylation site"  
 FT Modified-site 597  
 FT /note= "Potential phosphorylation site"  
 FT Modified-site 625  
 FT /note= "Potential phosphorylation site"  
 FT Region 633. .646  
 FT /label= Leucine-rich\_repeat  
 FT /note= "Signature sequence"  
 FT Modified-site 638  
 FT /note= "Potential phosphorylation site"  
 FT Modified-site 650  
 FT /note= "Potential phosphorylation site"  
 FT Region 656. .669  
 FT /label= Leucine-rich\_repeat  
 FT /note= "Signature sequence"  
 FT Region 659. .672  
 FT /label= Leucine-rich\_repeat  
 FT /note= "Signature sequence"  
 FT Modified-site 660  
 FT /note= "Potential phosphorylation site"  
 XX  
 PN WO200014248-A1.  
 XX  
 PD 16-MAR-2000.  
 XX  
 PF 03-SEP-1999; 99WO-US020287.  
 XX  
 PR 04-SEP-1998; 98US-00148904.  
 PR 04-SEP-1998; 98US-00155210.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Hillman JL, Yue H, Guegler KJ, Corley NC, Patterson C, Tang YT;  
 XX  
 DR WPI; 2000-256994/22.  
 DR N-PSDB; AAZ51682.  
 XX  
 PT Isolated cyclic nucleotide associated proteins useful for preventing,  
 PT diagnosing and treating cell proliferative, autoimmune/inflammatory,  
 PT neurological, vision, reproductive and smooth muscle disorders.  
 XX  
 PS Claim 1; Page 62-64; 78pp; English.  
 XX  
 CC The present sequence is a human cyclic nucleotide- associated protein-1  
 CC (CNAP-1), identified in Incyte clone 2267958, that is isolated from  
 CC UTRSNOT02 cDNA library. It is expressed in gastrointestinal,

CC reproductive, cardiovascular, haematopoietic/immune, nervous and urologic  
 CC tissues. CNAP sequences may be used for prevention, treatment and  
 CC diagnosis of diseases associated with altered CNAP expression such as,  
 CC cell proliferative disorders (e.g. arteriosclerosis, cirrhosis,  
 CC leukaemia, lymphoma and cancer of the breast, prostate, lung and brain),  
 CC autoimmune/inflammatory disorders (e.g. asthma, anaemia, diabetes  
 CC mellitus, multiple sclerosis and psoriasis), neurological disorders (e.g.  
 CC epilepsy, Alzheimer's/Parkinson's disease and strokes), vision disorders  
 CC (e.g. conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),  
 CC reproductive disorders (e.g. infertility, uterine fibroids, ectopic  
 CC pregnancies and impotence) and smooth muscle disorders (e.g. angina,  
 CC anaphylactic shock, Kearns-Sayre syndrome and hypertension). The coding  
 CC sequence can be used for gene therapy

XX

SQ Sequence 708 AA;

Quéry Match 82.9%; Score 34; DB 3; Length 708;  
 Best Local Similarity 77.8%; Pred. No. 5.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 :|||| |||  
 Db 337 MLSGLPDTV 345

# RESULT 13

ABU52881

ID ABU52881 standard; protein; 708 AA..

XX

AC ABU52881;

XX

DT 14-APR-2003 (first entry)

XX

DE Human uterus-derived protein from DKFZphutel\_19j11.

XX

KW Human; gene therapy; vaccine; disease treatment; detection.

XX

OS Homo sapiens.

XX

PN WO200112659-A2.

XX

PD 22-FEB-2001.

XX

PF 18-AUG-2000; 2000WO-IB001496.

XX

PR 18-AUG-1999; 99US-0149499P.

PR 28-SEP-1999; 99US-0156503P.

XX

PA (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX

PI Wiemann S;

XX

DR WPI; 2001-327840/34.

DR N-PSDB; ABX71314.

XX

PT Nucleic acids having the sequences of clones isolated from libraries of  
 PT different human tissues, useful in recombinant DNA methodologies.

XX  
PS Claim 21; Page 476; 1095pp; English.  
XX  
CC This invention describes novel polynucleotides and polypeptides isolated  
CC from human cDNA libraries which can be used for gene therapy or in  
CC vaccines. The polynucleotides of the invention and antibodies encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. The products of the  
CC invention may also be used to identify modulators of expression and  
CC activity and to down regulate expression and activity. The antibodies of  
CC the invention may also be used as diagnostic agents for detecting the  
CC presence of polypeptides in samples. This sequence represents a  
CC polypeptide described in the disclosure of the invention  
XX  
SQ Sequence 708 AA;

Query Match 82.9%; Score 34; DB 4; Length 708;  
Best Local Similarity 77.8%; Pred. No. 5.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV.9  
 :|||| |||  
Db 337 MLSGLPDTV 345

RESULT 14

ADN24197

ID ADN24197 standard; protein; 798 AA.

XX

AC ADN24197;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #6850.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.

XX

OS Bacteria.

XX

PN US2003233675-A1.

XX

PD 18-DEC-2003.

XX

PF 20-FEB-2003; 2003US-00369493.

XX

PR 21-FEB-2002; 2002US-0360039P.

XX

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX

DR WPI; 2004-061375/06.

XX

PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.

XX

PS Claim 1; SEQ ID NO 6850; 122pp; English.

XX

CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX

SQ Sequence 798 AA;

Query Match 82.9%; Score 34; DB 8; Length 798;

Best Local Similarity 77.8%; Pred. No. 6e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9

||: || ||

Db 75 LLNGLDDTV 83

RESULT 15

ADC15007

ID ADC15007 standard; protein; 803 AA.

XX

AC ADC15007;

XX

DT 18-DEC-2003 (first entry)

XX

DE Mouse adipocyte differentiation-associated protein SEQ ID NO:2.



XX  
KW mouse; anorectic; hypotensive; respiratory; antilipemic; antidiabetic;  
KW antiarteriosclerotic; cardiant; cerebroprotective; gene therapy; obesity;  
KW hypertension; hyperlipidaemia; diabetes; arteriosclerosis; cardiopathy;  
KW apoplexy; adipocyte differentiation.

XX  
OS Mus sp.

XX  
PN WO2003074694-A1.

XX  
PD 12-SEP-2003.

XX  
PF 25-FEB-2003; 2003WO-JP002019.

XX  
PR 25-FEB-2002; 2002JP-00048742.

XX  
PA (NISR ) NISSUI PHARM CO LTD.

XX  
PI Imagawa M, Oku Y;

XX  
DR WPI; 2003-712893/67.

DR N-PSDB; ADC15006.

XX  
PT Adipocyte differentiation-associated genes and their proteins, useful in  
PT diagnosis and developing drugs for lifestyle diseases such as obesity,  
PT hypertension, hyperlipemia and diabetes.

XX  
PS Claim 4; SEQ ID NO 2; 85pp; Japanese.

XX  
CC The invention relates to a novel adipocyte differentiation-associated  
CC polynucleotide as given in the specification, or its complementary  
CC strand. A protein of the invention has anorectic, hypotensive,  
CC respiratory, antilipemic, antidiabetic, antiarteriosclerotic, cardiant,  
CC and cerebroprotective activity. A polynucleotide of the invention may  
CC have a use in gene therapy. The genes and their encoded proteins are  
CC useful in diagnosis and developing drugs for lifestyle diseases such as  
CC obesity, hypertension, hyperlipidaemia, diabetes and arteriosclerosis-  
CC caused cardiopathy and apoplexy. The present sequence represents a  
CC protein of the invention.

XX  
SQ Sequence 803 AA;

Query Match 82.9%; Score 34; DB 7; Length 803;  
Best Local Similarity 77.8%; Pred. No. 6.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
:|||| |||  
Db 432 MLSGLPDTV 440

Search completed: February 10, 2005, 15:48:49  
Job time : 80.4648 secs

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 20.1549 Seconds  
(without alignments)  
33.334 Million cell updates/sec

Title: US-10-067-484-11  
Perfect score: 41  
Sequence: 1 LLSGLSDTV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	35	85.4	553	3	US-08-997-251-2	Sequence 2, Appli
2	34	82.9	2636	4	US-09-252-991A-25753	Sequence 25753, A
3	33	80.5	406	4	US-09-328-352-6564	Sequence 6564, Ap
4	33	80.5	518	4	US-09-252-991A-23491	Sequence 23491, A
5	33	80.5	1311	2	US-08-540-406-4	Sequence 4, Appli
6	33	80.5	1311	3	US-08-656-055-4	Sequence 4, Appli
7	33	80.5	1311	3	US-08-954-668-4	Sequence 4, Appli
8	33	80.5	1311	4	US-08-918-658-4	Sequence 4, Appli
9	33	80.5	1311	4	US-09-724-631-4	Sequence 4, Appli
10	33	80.5	1311	4	US-08-954-701A-4	Sequence 4, Appli
11	33	80.5	1311	5	PCT-US95-13233-4	Sequence 4, Appli
12	32	78.0	546	4	US-09-907-794A-250	Sequence 250, App
13	32	78.0	546	4	US-09-905-125A-250	Sequence 250, App
14	32	78.0	546	4	US-09-902-775A-250	Sequence 250, App
15	32	78.0	546	4	US-09-906-700-250	Sequence 250, App

16	32	78.0	546	4	US-09-903-603A-250	Sequence 250, App
17	32	78.0	546	4	US-09-904-920A-250	Sequence 250, App
18	32	78.0	546	4	US-09-909-064-250	Sequence 250, App
19	32	78.0	546	4	US-09-905-381A-250	Sequence 250, App
20	32	78.0	546	4	US-09-906-618-250	Sequence 250, App
21	32	78.0	553	3	US-08-997-251-4	Sequence 4, Appli
22	31	75.6	259	4	US-09-270-767-38568	Sequence 38568, A
23	31	75.6	259	4	US-09-270-767-53785	Sequence 53785, A
24	31	75.6	483	4	US-09-328-352-6271	Sequence 6271, Ap
25	30	73.2	161	4	US-10-101-464A-678	Sequence 678, App
26	30	73.2	185	4	US-09-270-767-35958	Sequence 35958, A
27	30	73.2	185	4	US-09-270-767-51175	Sequence 51175, A
28	30	73.2	266	4	US-09-248-796A-21194	Sequence 21194, A
29	30	73.2	304	4	US-09-902-540-12191	Sequence 12191, A
30	30	73.2	387	1	US-08-123-161A-12	Sequence 12, Appl
31	30	73.2	387	1	US-08-483-278-12	Sequence 12, Appl
32	30	73.2	392	4	US-09-949-016-9728	Sequence 9728, Ap
33	30	73.2	409	4	US-09-252-991A-18004	Sequence 18004, A
34	30	73.2	503	4	US-09-328-352-8228	Sequence 8228, Ap
35	30	73.2	793	4	US-09-489-039A-7449	Sequence 7449, Ap
36	30	73.2	966	4	US-09-902-540-14084	Sequence 14084, A
37	29	70.7	67	4	US-09-252-991A-25848	Sequence 25848, A
38	29	70.7	110	4	US-09-489-039A-10624	Sequence 10624, A
39	29	70.7	162	4	US-09-252-991A-24875	Sequence 24875, A
40	29	70.7	238	4	US-09-556-877-303	Sequence 303, App
41	29	70.7	238	4	US-09-620-412C-303	Sequence 303, App
42	29	70.7	238	4	US-09-598-419-303	Sequence 303, App
43	29	70.7	256	4	US-09-252-991A-30474	Sequence 30474, A
44	29	70.7	264	3	US-08-969-644-12	Sequence 12, Appl
45	29	70.7	264	3	US-08-444-189-12	Sequence 12, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-997-251-2

```
; Sequence 2, Application US/08997251
; Patent No. 6271440
; GENERAL INFORMATION:
;   APPLICANT: GUBLER, FRANZ J.
;   APPLICANT: JACOBSEN, JOHN V.
;   TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
;     STREET: 5370 Manhattan Boulevard
;     CITY: Boulder
;     STATE: CO
;     COUNTRY: US
;     ZIP: 80303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/997,251  
 ; FILING DATE: 23-DEC-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO AU96/00383  
 ; FILING DATE: 21-JUN-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: AU PN6470/95  
 ; FILING DATE: 09-NOV-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: AU PN3779/95  
 ; FILING DATE: 23-JUN-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Winner, Ellen P.  
 ; REGISTRATION NUMBER: 28,547  
 ; REFERENCE/DOCKET NUMBER: 110-97  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (303) 499-8080  
 ; TELEFAX: (303) 499-8089  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 553 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-997-251-2

Query Match 85.4%; Score 35; DB 3; Length 553;  
 Best Local Similarity 77.8%; Pred. No. 38;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
 || |||||:  
 Db 247 LLPGLSDTI 255

# RESULT 2

US-09-252-991A-25753  
 ; Sequence 25753, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 25753  
 ; LENGTH: 2636  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25753

Query Match 82.9%; Score 34; DB 4; Length 2636;  
Best Local Similarity 77.8%; Pred. No. 3.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||: || |||  
Db 371 LLAGLKDTV 379

RESULT 3

US-09-328-352-6564

; Sequence 6564, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6564  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii

US-09-328-352-6564

Query Match 80.5%; Score 33; DB 4; Length 406;  
Best Local Similarity 77.8%; Pred. No. 71;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||: || |||  
Db 77 LLAGLPDTV 85

RESULT 4

US-09-252-991A-23491

; Sequence 23491, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23491

; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23491

Query Match 80.5%; Score 33; DB 4; Length 518;  
Best Local Similarity 77.8%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| || ||  
Db 177 LLDGLDDTV 185

RESULT 5

US-08-540-406-4

; Sequence 4, Application US/08540406

; Patent No. 5837538

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; APPLICANT: GOODRICH, LISA V

; APPLICANT: JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/540,406

; FILING DATE: 06-OCT-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1311 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-540-406-4

Query Match 80.5%; Score 33; DB 2; Length 1311;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
      |||||  
Db 901 LSGLSDT 907

RESULT 6

US-08-656-055-4

; Sequence 4, Application US/08656055

; Patent No. 6027882

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; APPLICANT: GOODRICH, LISA V

; APPLICANT: JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/656,055

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/540,406

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1311 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-656-055-4

Query Match 80.5%; Score 33; DB 3; Length 1311;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
|||||||  
Db 901 LSGLSDT 907

RESULT 7

US-08-954-668-4

; Sequence 4, Application US/08954668  
; Patent No. 6172200  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; APPLICANT: GOODRICH, LISA V  
; APPLICANT: JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/954,668  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36709  
; REFERENCE/DOCKET NUMBER: SUV-003.06  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1311 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-954-668-4

Query Match 80.5%; Score 33; DB 3; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
|||||||  
Db 901 LSGLSDT 907

RESULT 8



US-08-918-658-4  
; Sequence 4, Application US/08918658  
; Patent No. 6429354  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; GOODRICH, LISA V  
; JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/918,658  
; FILING DATE: 22-Aug-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/656,055  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/540,406  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: a60190-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1311 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-918-658-4

Query Match 80.5%; Score 33; DB 4; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
|||||  
Db 901 LSGLSDT 907

RESULT 9

US-09-724-631-4  
; Sequence 4, Application US/09724631  
; Patent No. 6551782  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; GOODRICH, LISA V  
; JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/724,631  
; FILING DATE: 28-No. 6551782-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/656,055  
; FILING DATE: 1996-05-31  
; APPLICATION NUMBER: 08/540,406  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: a60190-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1311 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-724-631-4

Query Match 80.5%; Score 33; DB 4; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
|||||  
Db 901 LSGLSDT 907

RESULT 10

US-08-954-701A-4  
; Sequence 4, Application US/08954701A  
; Patent No. 6610507  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; APPLICANT: GOODRICH, LISA V  
; APPLICANT: JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/954,701A  
; FILING DATE: 20-OCT-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36709  
; REFERENCE/DOCKET NUMBER: SUV-003.08  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1311 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-954-701A-4

Query Match 80.5%; Score 33; DB 4; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
|||||  
Db 901 LSGLSDT 907

RESULT 11

PCT-US95-13233-4

; Sequence 4, Application PC/TUS9513233  
; GENERAL INFORMATION:  
; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
 ; STREET: Four Embarcadero Center, Suite 3400  
 ; CITY: San Francisco  
 ; STATE: CA  
 ; COUNTRY: US  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/13233  
 ; FILING DATE: 06-OCT-1990  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Rowland, Bertram I  
 ; REGISTRATION NUMBER: 20015  
 ; REFERENCE/DOCKET NUMBER: a60190-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-781-1989  
 ; TELEFAX: 415-398-3249  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1311 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 PCT-US95-13233-4

Query Match 80.5%; Score 33; DB 5; Length 1311;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDT 8  
 |||||  
 Db 901 LSGLSDT 907

RESULT 12  
 US-09-907-794A-250  
 ; Sequence 250, Application US/09907794A  
 ; Patent No. 6635468  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genentech, Inc.  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,794A  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 250  
; LENGTH: 546  
; TYPE: PRT

; ORGANISM: Homo Sapien  
US-09-907-794A-250

Query Match 78.0%; Score 32; DB 4; Length 546;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
: |||: |||  
Db 170 MLSGIPDTV 178

RESULT 13

US-09-905-125A-250

; Sequence 250, Application US/09905125A  
; Patent No. 6664376

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/905,125A

; CURRENT FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944  
 ; PRIOR FILING DATE: 1999-09-13  
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090  
 ; PRIOR FILING DATE: 1999-09-15  
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547  
 ; PRIOR FILING DATE: 1999-09-15  
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214  
 ; PRIOR FILING DATE: 1999-11-29  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313  
 ; PRIOR FILING DATE: 1999-11-30  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911  
 ; PRIOR FILING DATE: 1999-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999  
 ; PRIOR FILING DATE: 1999-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219  
 ; PRIOR FILING DATE: 2000-01-05  
 ; NUMBER OF SEQ ID NOS: 423  
 ; SEQ ID NO 250  
 ; LENGTH: 546  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-09-905-125A-250

Query Match 78.0%; Score 32; DB 4; Length 546;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
 :|||: |||  
 Db 170 MLSGIPDTV 178

# RESULT 14

US-09-902-775A-250  
 ; Sequence 250, Application US/09902775A  
 ; Patent No. 6686451  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genentech, Inc.  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/902,775A  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 250  
; LENGTH: 546  
; TYPE: PRT  
; ORGANISM: Homo Sapien



US-09-902-775A-250

Query Match 78.0%; Score 32; DB 4; Length 546;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
:||||: |||  
Db 170 MLSGIPDTV 178

RESULT 15

US-09-906-700-250

; Sequence 250, Application US/09906700  
; Patent No. 6723535  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/906,700  
; CURRENT FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13  
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090  
 ; PRIOR FILING DATE: 1999-09-15  
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547  
 ; PRIOR FILING DATE: 1999-09-15  
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214  
 ; PRIOR FILING DATE: 1999-11-29  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313  
 ; PRIOR FILING DATE: 1999-11-30  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565  
 ; PRIOR FILING DATE: 1999-12-02  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911  
 ; PRIOR FILING DATE: 1999-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999  
 ; PRIOR FILING DATE: 1999-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219  
 ; PRIOR FILING DATE: 2000-01-05  
 ; NUMBER OF SEQ ID NOS: 423  
 ; SEQ ID NO 250  
 ; LENGTH: 546  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-09-906-700-250

Query Match 78.0%; Score 32; DB 4; Length 546;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
 :|||: |||  
 Db 170 MLSGIPDTV 178

Search completed: February 10, 2005, 16:02:11  
 Job time : 20.1549 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:49:10 ; Search time 53.8732 Seconds  
 (without alignments)  
 54.586 Million cell updates/sec

Title: US-10-067-484-11  
 Perfect score: 41  
 Sequence: 1 LLSGLSDTV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Match	Length			
1	41	100.0	9	14	US-10-067-484-11	Sequence 11, Appl
2	41	100.0	9	14	US-10-067-620-11	Sequence 11, Appl
3	36	87.8	100	15	US-10-424-599-223261	Sequence 223261,
4	35	85.4	122	9	US-09-738-626-3620	Sequence 3620, Ap
5	35	85.4	553	15	US-10-389-566-1119	Sequence 1119, Ap
6	35	85.4	553	15	US-10-389-566-2355	Sequence 2355, Ap
7	34	82.9	708	15	US-10-120-801-103	Sequence 103, App
8	34	82.9	708	15	US-10-120-801-104	Sequence 104, App
9	34	82.9	798	15	US-10-369-493-6850	Sequence 6850, Ap
10	34	82.9	5149	15	US-10-282-122A-43633	Sequence 43633, A
11	33	80.5	69	15	US-10-424-599-178445	Sequence 178445,
12	33	80.5	76	15	US-10-282-122A-77386	Sequence 77386, A
13	33	80.5	113	15	US-10-369-493-18817	Sequence 18817, A

14	33	80.5	692	16	US-10-437-963-144927	Sequence 144927,
15	33	80.5	845	16	US-10-467-595-21	Sequence 21, Appl
16	33	80.5	1311	8	US-08-954-701A-4	Sequence 4, Appli
17	33	80.5	1311	10	US-09-754-032-4	Sequence 4, Appli
18	33	80.5	1311	14	US-10-421-446-4	Sequence 4, Appli
19	32	78.0	238	14	US-10-156-761-10224	Sequence 10224, A
20	32	78.0	300	9	US-09-925-299-915	Sequence 915, App
21	32	78.0	300	10	US-09-925-299-915	Sequence 915, App
22	32	78.0	338	10	US-09-908-006A-53	Sequence 53, Appl
23	32	78.0	397	9	US-09-815-242-12461	Sequence 12461, A
24	32	78.0	423	16	US-10-437-963-128172	Sequence 128172,
25	32	78.0	546	9	US-09-909-320-250	Sequence 250, App
26	32	78.0	546	9	US-09-909-088B-250	Sequence 250, App
27	32	78.0	546	9	US-09-905-291A-250	Sequence 250, App
28	32	78.0	546	9	US-09-902-853-250	Sequence 250, App
29	32	78.0	546	9	US-09-907-824-250	Sequence 250, App
30	32	78.0	546	9	US-09-907-841-250	Sequence 250, App
31	32	78.0	546	10	US-09-904-011-250	Sequence 250, App
32	32	78.0	546	10	US-09-903-640-250	Sequence 250, App
33	32	78.0	546	10	US-09-908-093-250	Sequence 250, App
34	32	78.0	546	10	US-09-906-742-250	Sequence 250, App
35	32	78.0	546	10	US-09-906-838-250	Sequence 250, App
36	32	78.0	546	10	US-09-907-613-250	Sequence 250, App
37	32	78.0	546	10	US-09-907-942-250	Sequence 250, App
38	32	78.0	546	10	US-09-904-859-250	Sequence 250, App
39	32	78.0	546	10	US-09-909-204-250	Sequence 250, App
40	32	78.0	546	10	US-09-904-820-250	Sequence 250, App
41	32	78.0	546	10	US-09-904-786-250	Sequence 250, App
42	32	78.0	546	10	US-09-906-646-250	Sequence 250, App
43	32	78.0	546	10	US-09-906-700-250	Sequence 250, App
44	32	78.0	546	10	US-09-903-786-250	Sequence 250, App
45	32	78.0	546	10	US-09-902-903-250	Sequence 250, App

#### ALIGNMENTS

#### RESULT 1

US-10-067-484-11

```
; Sequence 11, Application US/10067484
; Publication No. US20030170763A1
; GENERAL INFORMATION:
; APPLICANT: Buchanan, Bob B.
; APPLICANT: del Val, Gregorio
; APPLICANT: Frick, Oscar L.
; TITLE OF INVENTION: RAGWEED ALLERGENS
; FILE REFERENCE: 416272000200
; CURRENT APPLICATION NUMBER: US/10/067,484
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/266,686
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Ragweed
```

US-10-067-484-11

Query Match 100.0%; Score 41; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
      |||||  
Db 1 LLSGLSDTV 9

RESULT 2

US-10-067-620-11

; Sequence 11, Application US/10067620  
; Publication No. US20030180225A1  
; GENERAL INFORMATION:  
; APPLICANT: Buchanan, Bob B.  
; APPLICANT: del Val, Gregorio  
; APPLICANT: Frick, Oscar L.  
; APPLICANT: Teuber, Suzanne S.  
; TITLE OF INVENTION: WALNUT AND RYEGRASS ALLERGENS  
; FILE REFERENCE: 416272003400  
; CURRENT APPLICATION NUMBER: US/10/067,620  
; CURRENT FILING DATE: 2002-02-04  
; PRIOR APPLICATION NUMBER: US 60/266,686  
; PRIOR FILING DATE: 2001-02-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Ragweed  
US-10-067-620-11

Query Match 100.0%; Score 41; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
      |||||  
Db 1 LLSGLSDTV 9

RESULT 3

US-10-424-599-223261

; Sequence 223261, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 223261  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(100)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_43635C.1.pep  
US-10-424-599-223261

Query Match 87.8%; Score 36; DB 15; Length 100;  
Best Local Similarity 88.9%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||| |||  
Db 70 LLSGLRDTV 78

#### RESULT 4

US-09-738-626-3620  
; Sequence 3620, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3620  
; LENGTH: 122  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3620

Query Match 85.4%; Score 35; DB 9; Length 122;

Best Local Similarity 87.5%; Pred. No. 24;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
:|||||  
Db 52 ISGLSDTV 59

RESULT 5

US-10-389-566-1119

; Sequence 1119, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1119  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Hordeum vulgare  
US-10-389-566-1119

Query Match 85.4%; Score 35; DB 15; Length 553;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 247 LLPGLSDTI 255

RESULT 6

US-10-389-566-2355

; Sequence 2355, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25

; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2355  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Hordeum vulgare  
US-10-389-566-2355

Query Match 85.4%; Score 35; DB 15; Length 553;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
|| |||||:  
Db 247 LLPGLSDTI 255

RESULT 7

US-10-120-801-103  
; Sequence 103, Application US/10120801  
; Publication No. US20030203843A1  
; GENERAL INFORMATION:  
; APPLICANT: Pena, Carol  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Mehraban, Fuad  
; APPLICANT: Topper, James N.  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Wasserman, Scott  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Gunther, Erik  
; APPLICANT: Komuves, Laszlo  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-340  
; CURRENT APPLICATION NUMBER: US/10/120,801  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 60/285748  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: 60/286068  
; PRIOR FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: 60/286292  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: 60/288334  
; PRIOR FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 60/291241  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 60/322284  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/285609  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 155



; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 103  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: human  
US-10-120-801-103

Query Match 82.9%; Score 34; DB 15; Length 708;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
:|||||  
Db 337 MLSGLPDTV 345

RESULT 8

US-10-120-801-104  
; Sequence 104, Application US/10120801  
; Publication No. US20030203843A1  
; GENERAL INFORMATION:  
; APPLICANT: Pena, Carol  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Mehraban, Fuad  
; APPLICANT: Topper, James N.  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Wasserman, Scott  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Gunther, Erik  
; APPLICANT: Komuves, Laszlo  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-340  
; CURRENT APPLICATION NUMBER: US/10/120,801  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 60/285748  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: 60/286068  
; PRIOR FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: 60/286292  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: 60/288334  
; PRIOR FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 60/291241  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 60/322284  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/285609  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 104  
; LENGTH: 708

; TYPE: PRT  
; ORGANISM: mouse  
US-10-120-801-104

Query Match 82.9%; Score 34; DB 15; Length 708;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
:|||| |||  
Db 337 MLSGLPDTV 345

RESULT 9

US-10-369-493-6850

; Sequence 6850, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION  
OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 6850  
; LENGTH: 798  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-369-493-6850

Query Match 82.9%; Score 34; DB 15; Length 798;  
Best Local Similarity 77.8%; Pred. No. 3.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||:|| |||  
Db 75 LLNGLDDTV 83

RESULT 10

US-10-282-122A-43633

; Sequence 43633, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari

```

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43633
; LENGTH: 5149
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43633

```

```

Query Match      82.9%; Score 34; DB 15; Length 5149;
Best Local Similarity 77.8%; Pred. No. 2.6e+03;
Matches      7; Conservative      1; Mismatches      1; Indels      0; Gaps      0;

```

```

Qy      1 LLSGLSDTV 9
      ||:|| |||
Db      357 LLAGLKDTV 365

```

```

RESULT 11
US-10-424-599-178445
; Sequence 178445, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

```

; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
 With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 178445  
 ; LENGTH: 69  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_132153C.1.pep  
 US-10-424-599-178445

Query Match 80.5%; Score 33; DB 15; Length 69;  
 Best Local Similarity 87.5%; Pred. No. 32;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
 | | | | |  
 Db 3 LDGLSDTV 10

# RESULT 12

US-10-282-122A-77386  
 ; Sequence 77386, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 77386  
 ; LENGTH: 76  
 ; TYPE: PRT  
 ; ORGANISM: Vibrio cholerae  
 US-10-282-122A-77386

Query Match 80.5%; Score 33; DB 15; Length 76;  
 Best Local Similarity 87.5%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
 | |||||  
 Db 17 LDGLSDTV 24

# RESULT 13

US-10-369-493-18817  
 ; Sequence 18817, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION  
 OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 18817  
 ; LENGTH: 113  
 ; TYPE: PRT  
 ; ORGANISM: Anabaena PCC7120  
 US-10-369-493-18817

Query Match 80.5%; Score 33; DB 15; Length 113;  
 Best Local Similarity 77.8%; Pred. No. 56;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9

Db                    || ||||:|  
35 LLKGLSDSV 43

RESULT 14

US-10-437-963-144927  
; Sequence 144927, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules  
Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 144927  
; LENGTH: 692  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_45696C.1.pep  
US-10-437-963-144927

Query Match                    80.5%;    Score 33;    DB 16;    Length 692;  
Best Local Similarity    77.8%;    Pred. No. 4.4e+02;  
Matches        7;    Conservative        1;    Mismatches        1;    Indels        0;    Gaps        0;

Qy                    1 LLSGLSDTV 9  
                      || .||:|||  
Db                    283 LLPGLNDTV 291

RESULT 15

US-10-467-595-21  
; Sequence 21, Application US/10467595  
; Publication No. US20040166501A1  
; GENERAL INFORMATION:  
; APPLICANT: AZIMZAI, Yalda; YUE, Henry;  
; APPLICANT: DING, Li; NGUYEN, Danniel B.;  
; APPLICANT: GANDHI, Ameena R.; BURFORD, Neil;  
; APPLICANT: THANGAVELU, Kavitha; ELLIOTT, Vicki S.;  
; APPLICANT: RAMKUMAR, Jayalaxmi; YAO, Monique G.;  
; APPLICANT: LAL, Preeti G.; TANG, Y. Tom;  
; APPLICANT: SWARNAKAR, Anita; WARREN, Bridget A.;  
; APPLICANT: WALIA, Narinder K.; POLICKY, Jennifer L.;  
; APPLICANT: XU, Yunming; HONCHELL, Cynthia D.;  
; APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.;  
; APPLICANT: DUGGAN, Brendan M.; LU, Dyung Aina M.;

; APPLICANT: GIETZEN, Kimberly J.; JACKSON, Jennifer L.;  
 ; APPLICANT: RAUMANN, Bridget E.; LU, Yan;  
 ; APPLICANT: KAREHT, Stephanie K.; TRAN, Uyen K.;  
 ; APPLICANT: RICHARDSON, Thomas W.; EMERLING, Brook M.;  
 ; APPLICANT: HAFALIA April J.A.; BURRILL, John D.;  
 ; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;  
 ; APPLICANT: KABLE, Amy E.; GORVAD, Ann E.  
 ; TITLE OF INVENTION: RECEPTORS AND MEMBRANE-ASSOCIATED PROTEINS  
 ; FILE REFERENCE: PF-0894 PCT  
 ; CURRENT APPLICATION NUMBER: US/10/467,595  
 ; CURRENT FILING DATE: 2003-08-06  
 ; PRIOR APPLICATION NUMBER: PCT/US02/03868  
 ; PRIOR FILING DATE: 2002-02-05  
 ; PRIOR APPLICATION NUMBER: US 60/267,201  
 ; PRIOR FILING DATE: 2001-02-06  
 ; PRIOR APPLICATION NUMBER: US 60/269,580  
 ; PRIOR FILING DATE: 2001-02-16  
 ; PRIOR APPLICATION NUMBER: US 60/282,679  
 ; PRIOR FILING DATE: 2001-04-09  
 ; PRIOR APPLICATION NUMBER: US 60/288,295  
 ; PRIOR FILING DATE: 2001-05-02  
 ; PRIOR APPLICATION NUMBER: US 60/348,687  
 ; PRIOR FILING DATE: 2002-01-14  
 ; NUMBER OF SEQ ID NOS: 86  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 21  
 ; LENGTH: 845  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No: 4973984CD1  
 US-10-467-595-21

Query Match 80.5%; Score 33; DB 16; Length 845;  
 Best Local Similarity 75.0%; Pred. No. 5.5e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
 |:|||||:  
 Db 267 LNGLSDTI 274

Search completed: February 10, 2005, 16:41:35  
 Job time : 54.8732 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 13.9437 Seconds  
 (without alignments)  
 62.104 Million cell updates/sec

Title: US-10-067-484-11

Perfect score: 41  
Sequence: 1 LLSGLSDTV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	%		Query Match Length DB	ID	Description
	Score	Match			
1	38	92.7	229 2	T22037	hypothetical prote
2	37	90.2	295 2	T24036	hypothetical prote
3	36	87.8	75 2	A86487	unknown protein [i
4	35	85.4	397 2	T46710	multidrug resistan
5	35	85.4	397 2	AE1613	multidrug-efflux t
6	35	85.4	397 2	AI1250	multidrug-efflux t
7	35	85.4	553 2	T06179	myb-related protei
8	34	82.9	381 2	B84935	acetylornithine de
9	34	82.9	798 2	T28906	hypothetical prote
10	34	82.9	5149 2	F83345	probable non-ribos
11	33	80.5	76 2	C82122	ferrous iron trans
12	33	80.5	324 2	S28309	hypothetical prote
13	33	80.5	465 2	AB3536	penicillin amidase
14	32	78.0	238 2	T29082	probable transcrip
15	32	78.0	315 2	A70137	smg protein homolo
16	32	78.0	318 2	H71406	probable reverse t
17	32	78.0	373 2	AD0030	conserved hypothet
18	32	78.0	394 2	B89786	hypothetical prote
19	32	78.0	467 2	A82664	conserved hypothet
20	32	78.0	498 2	B90456	hypothetical prote
21	32	78.0	499 2	G98180	lysyl-tRNA synthet
22	32	78.0	499 2	AE3106	lysyl-tRNA synthet
23	32	78.0	553 2	T03762	myb-related transc
24	32	78.0	726 2	A90771	hypothetical prote
25	32	78.0	726 2	E85633	hypothetical prote



26	32	78.0	726	2	C64839	probable ATPase yc
27	32	78.0	1249	2	T47885	beta-tubulin cofac
28	31	75.6	100	2	S77030	hypothetical prote
29	31	75.6	130	2	S58829	hypothetical prote
30	31	75.6	179	2	AF2242	hypothetical prote
31	31	75.6	297	2	T51273	promoter-binding f
32	31	75.6	340	2	G82047	conserved hypothet
33	31	75.6	364	2	B85430	hypothetical prote
34	31	75.6	378	2	C82049	acetylornithine de
35	31	75.6	427	2	AH2350	hypothetical prote
36	31	75.6	445	2	S64063	hypothetical prote
37	31	75.6	484	2	AG2820	two component sens
38	31	75.6	497	2	G97598	probable two-compo
39	31	75.6	573	2	AB2401	hypothetical prote
40	31	75.6	730	2	T43317	pgl-1 protein - Ca
41	31	75.6	771	2	T29177	hypothetical prote
42	31	75.6	1444	1	A30588	140K adhesin precu
43	30	73.2	190	2	T51590	membrane protein,
44	30	73.2	214	2	B33985	wound-inducible ch
45	30	73.2	295	2	AG0848	probable AraC-fami

#### ALIGNMENTS

##### RESULT 1

T22037

hypothetical protein F40F8.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T22037

R;MacDougall, R.

submitted to the EMBL Data Library, February 1996

A;Reference number: Z19505

A;Accession: T22037

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-229 <WIL>

A;Cross-references: UNIPROT:Q20231; EMBL:Z69302; PIDN:CAA93265.1; GSPDB:GN00020;

CESP:F40F8.3

A;Experimental source: clone F40F8

C;Genetics:

A;Gene: CESP:F40F8.3

A;Map position: 2

A;Introns: 25/1; 61/3; 105/3; 126/1

C;Superfamily: Caenorhabditis elegans hypothetical protein F40F8.3

Query Match 92.7%; Score 38; DB 2; Length 229;

Best Local Similarity 77.8%; Pred. No. 2.4;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9

:|||||:

Db 209 ILSGLSDTI 217

##### RESULT 2

T24036

hypothetical protein R07H5.3 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T24036

R;Lennard, N.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19833

A;Accession: T24036

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-295 <WIL>

A;Cross-references: UNIPROT:Q93931; EMBL:Z81107; PIDN:CAB03226.1; GSPDB:GN00022; CESP:R07H5.3

A;Experimental source: clone R07H5

C;Genetics:

A;Gene: CESP:R07H5.3

A;Map position: 4

A;Introns: 61/1; 113/2; 165/2; 186/3; 232/3; 270/3

Query Match 90.2%; Score 37; DB 2; Length 295;

Best Local Similarity 77.8%; Pred. No. 5.1;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9

||:||||:

Db 238 LLTGLSDTI 246

RESULT 3

A86487

unknown protein [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: A86487

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86487

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-75 <STO>

A;Cross-references: UNIPROT:Q9C8F3; GB:AE005172; NID:g11038490; PIDN:AAG27768.1;  
GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match 87.8%; Score 36; DB 2; Length 75;  
Best Local Similarity 88.9%; Pred. No. 1.8;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||| |||: ||  
Db 15 LLSGLSNTV 23

#### RESULT 4

T46710

multidrug resistance transporter [imported] - *Listeria monocytogenes*

C;Species: *Listeria monocytogenes*

C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004

C;Accession: T46710

R;Huillet, E.E.H.; Larpin, S.; Pardon, P.; Berche, P.

FEMS Microbiol. Lett. 174, 265-272, 1999

A;Title: Identification of a new locus in *Listeria monocytogenes* involved in  
cellobiose-dependent repression of hly expression.

A;Reference number: Z23136; MUID:99271176; PMID:10339818

A;Accession: T46710

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-397 <HUI>

A;Cross-references: UNIPROT:Q9X769; EMBL:AJ009627; NID:g4914621;

PIDN:CAB43715.1; PID:g4914624

C;Genetics:

A;Gene: lltB

Query Match 85.4%; Score 35; DB 2; Length 397;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 89 LLCGLSDTI 97

#### RESULT 5

AE1613

multidrug-efflux transporter homolog lin1446 [imported] - *Listeria innocua*  
(strain Clip11262)

C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AE1613

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;

Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,

E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.;

Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.;

Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.;

Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AE1613

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-397 <GLA>

A;Cross-references: UNIPROT:Q92BU8; GB:AL592022; PIDN:CAC96677.1; PID:g16413919; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin1446

Query Match 85.4%; Score 35; DB 2; Length 397;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 89 LLCGLSDTI 97

#### RESULT 6

AI1250

multidrug-efflux transporter homolog lmo1409 [imported] - *Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AI1250

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AI1250

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-397 <GLA>

A;Cross-references: UNIPROT:Q9X769; GB:NC\_003210; PIDN:CAC99487.1; PID:g16410838; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo1409

Query Match 85.4%; Score 35; DB 2; Length 397;

Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 89 LLCGLSDTI 97

RESULT 7

T06179

myb-related protein - barley

C;Species: Hordeum vulgare (barley)

C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T06179

R;Gubler, F.; Kalla, R.; Roberts, J.K.; Jacobsen, J.V.

Plant Cell 7, 1879-1891, 1995

A;Title: Gibberellin regulated expression of a Myb gene in barley aleurone cells: evidence for Myb transactivation of a high-pI alpha-amylase gene promoter.

A;Reference number: Z15511; MUID:96093427; PMID:8535141

A;Accession: T06179

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-553 <GUB>

A;Cross-references: UNIPROT:Q96464; EMBL:X87690; NID:g1200238; PIDN:CAA61021.1; PID:g1200239

A;Experimental source: cv. Himalaya

C;Genetics:

A;Gene: Gam1

C;Function:

A;Description: transcriptional activator of alpha-amylase gene promoter

C;Superfamily: petunia myb-related protein 3; myb DNA-binding repeat homology

F;39-91/Domain: myb DNA-binding repeat homology <MYB1>

F;92-142/Domain: myb DNA-binding repeat homology <MYB>

Query Match 85.4%; Score 35; DB 2; Length 553;  
Best Local Similarity 77.8%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 247 LLPGLSDTI 255

RESULT 8

B84935

acetylornithine deacetylase (EC 3.5.1.16) [imported] - Buchnera sp. (strain APS)

C;Species: Buchnera sp.

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002

C;Accession: B84935

R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.

A;Reference number: A84930; MUID:20445173; PMID:10993077

A;Accession: B84935

A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-381 <STO>  
A;Cross-references: GB:AP000398; GSPDB:GN00144  
A;Experimental source: strain APS  
C;Genetics:  
A;Gene: argE; BU047  
C;Superfamily: succinyl-di-aminopimelate desuccinylase  
C;Keywords: hydrolase

Query Match 82.9%; Score 34; DB 2; Length 381;  
Best Local Similarity 88.9%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||| |||  
Db 75 LLSGHSDTV 83

RESULT 9

T28906

hypothetical protein T25B6.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T28906

R;Favello; A.

submitted to the EMBL Data Library, December 1995

A;Description: The sequence of C. elegans cosmid T25B6.

A;Reference number: Z20541

A;Accession: T28906

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-798 <FAV>

A;Cross-references: UNIPROT:Q22763; EMBL:U41546; PIDN:AAC48223.1; GSPDB:GN00028;  
CESP:T25B6.2

A;Experimental source: strain Bristol N2; clone T25B6

C;Genetics:

A;Gene: CESP:T25B6.2

A;Map position: X

A;Introns: 11/1; 96/3; 148/3; 174/2; 228/3; 299/2; 343/3; 411/2; 471/3; 492/1;  
749/3

C;Superfamily: neprilysin

Query Match 82.9%; Score 34; DB 2; Length 798;  
Best Local Similarity 77.8%; Pred. No. 67;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||: || |||  
Db 75 LLNGLDDTV 83

RESULT 10

F83345

probable non-ribosomal peptide synthetase PA2402 [imported] - Pseudomonas  
aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C;Accession: F83345  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: F83345  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-5149 <STO>  
A;Cross-references: UNIPROT:Q9I179; GB:AE004667; GB:AE004091; NID:g9948444; PIDN:AAG05790.1; GSPDB:GN00131; PASP:PA2402  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA2402  
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein  
F;557-998/Domain: acetate-CoA ligase homology <ACL1>  
F;1015-1082/Domain: acyl carrier protein homology <ACP1>  
F;2034-2479/Domain: acetate-CoA ligase homology <ACL2>  
F;2496-2564/Domain: acyl carrier protein homology <ACP2>  
F;3084-3525/Domain: acetate-CoA ligase homology <ACL3>  
F;3541-3608/Domain: acyl carrier protein homology <ACP3>  
F;4583-5040/Domain: acetate-CoA ligase homology <ACL4>  
F;5057-5125/Domain: acyl carrier protein homology <ACP4>  
F;1047,2528,3573,5089/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 82.9%; Score 34; DB 2; Length 5149;  
Best Local Similarity 77.8%; Pred. No. 5.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
||: || |||  
Db 357 LLAGLKDTV 365

RESULT 11  
C82122  
ferrous iron transport protein A VC2078 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
C;Species: *Vibrio cholerae*  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: C82122  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: C82122  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-76 <HEI>  
A;Cross-references: UNIPROT:Q9KQC2; GB:AE004281; GB:AE003852; NID:g9656616;  
PIDN:AAF95224.1; GSPDB:GN00126; TIGR:VC2078  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC2078  
A;Map position: 1

Query Match 80.5%; Score 33; DB 2; Length 76;  
Best Local Similarity 87.5%; Pred. No. 8;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
| | | | |  
Db 17 LDGLSDTV 24

RESULT 12

S28309

hypothetical protein F02A9.1b - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Jul-2004

C;Accession: S28309

R;Berks, M.

submitted to the EMBL Data Library, December 1992

A;Reference number: S28296

A;Accession: S28309

A;Molecule type: DNA

A;Residues: 1-324 <BER>

A;Cross-references: UNIPROT:P34381; EMBL:Z19555; NID:g6705; PID:g6707

C;Genetics:

A;Introns: 49/3; 133/2; 245/2

Query Match 80.5%; Score 33; DB 2; Length 324;  
Best Local Similarity 55.6%; Pred. No. 40;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
| : | | : | | :  
Db 278 LISGIADTI 286

RESULT 13

AB3536

penicillin amidase (EC 3.5.1.11) [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004

C;Accession: AB3536

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.;

Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.;

Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.;

Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn,

R.; Kyrpides, N.; Overbeek, R.



Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.

A;Reference number: AD3252; PMID:11756688

A;Accession: AB3536

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-465 <KUR>

A;Cross-references: UNIPROT:Q8YDG4; GB:AE008918; PIDN:AAL53453.1; PID:g17984352; GSPDB:GN00191

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEII0212

A;Map position: II

C;Keywords: hydrolase

Query Match 80.5%; Score 33; DB 2; Length 465;  
Best Local Similarity 87.5%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
| | | | |  
Db 16 LQGLSDTV 23

#### RESULT 14

T29082

probable transcription regulator - *Streptomyces coelicolor*

C;Species: *Streptomyces coelicolor*

C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T29082

R;Redenbach, M.; Kieser, H.M.; Denapaite, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopwood, D.A.

Mol. Microbiol. 21, 77-96, 1996

A;Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome.

A;Reference number: Z20556; MUID:97000351; PMID:8843436

A;Accession: T29082

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-238 <RED>

A;Cross-references: UNIPROT:O86533; EMBL:AL031124; PIDN:CAA20000.1

C;Genetics:

A;Note: SC1C2.33c

C;Superfamily: acetate operon repressor

Query Match 78.0%; Score 32; DB 2; Length 238;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LSGLSDTV 9  
| | | | |  
Db 118 LSGLRDTV 125

#### RESULT 15

A70137

smg protein homolog - Lyme disease spirochete

C;Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 09-Jul-2004

C;Accession: A70137

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman, J.; Utterback, T.; Watthey, L.; McDonald, L.; Artiach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: A70137

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-315 <KLE>

A;Cross-references: UNIPROT:Q44773; GB:AE001137; GB:AE000783; NID:g2688160;

PIDN:AAC66651.1; PID:g2688169; TIGR:BB0297

A;Experimental source: strain B31

Query Match 78.0%; Score 32; DB 2; Length 315;  
Best Local Similarity 66.7%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|:||||| :  
Db 212 LVSGLSDAI 220

Search completed: February 10, 2005, 15:59:38

Job time : 14.9437 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2005, 15:38:08 ; Search time 65.662 Seconds  
(without alignments)  
70.188 Million cell updates/sec

Title: US-10-067-484-11  
Perfect score: 41  
Sequence: 1 LLSGLSDTV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	38	92.7	229	2	Q20231	Q20231	caenorhabdi
2	38	92.7	394	2	Q8NYJ2	Q8nyj2	staphylococ
3	38	92.7	394	2	Q6GCP4	Q6gcp4	staphylococ
4	38	92.7	635	2	Q7RHZ2	Q7rhz2	plasmodium
5	37	90.2	270	2	Q93931	Q93931	caenorhabdi
6	37	90.2	295	2	Q7JL05	Q7jl05	caenorhabdi
7	36	87.8	75	2	Q9C8F3	Q9c8f3	arabidopsis
8	36	87.8	548	2	Q9XEP0	Q9xep0	lolium temu
9	35	85.4	122	2	Q8NU44	Q8nu44	corynebacte
10	35	85.4	312	2	Q8VFH5	Q8vfh5	mus musculu
11	35	85.4	312	2	Q7TRI4	Q7tri4	mus musculu
12	35	85.4	391	2	Q71ZR1	Q71zr1	listeria mo
13	35	85.4	392	2	Q9F427	Q9f427	listeria mo
14	35	85.4	397	2	Q92BU8	Q92bu8	listeria in
15	35	85.4	397	2	Q9X769	Q9x769	listeria mo
16	35	85.4	471	2	Q6KGM2	Q6kgm2	bacteriopha
17	35	85.4	503	2	Q6IX76	Q6ix76	triticum ae
18	35	85.4	546	2	Q9XFN2	Q9xfn2	avena sativ
19	35	85.4	553	2	Q96464	Q96464	hordeum vul
20	35	85.4	553	2	Q9ARM1	Q9arm1	hordeum vul
21	34	82.9	92	2	Q71IU7	Q71iu7	lactobacill
22	34	82.9	278	2	Q9WW90	Q9ww90	lactococcus
23	34	82.9	281	2	Q7QS90	Q7qs90	giardia lam
24	34	82.9	329	2	Q8KN25	Q8kn25	comamonas a
25	34	82.9	336	2	Q8RL40	Q8rl40	comamonas a
26	34	82.9	381	1	ARGE_BUCAI	P57155	buchnera ap
27	34	82.9	708	2	Q8R3G5	Q8r3g5	mus musculu
28	34	82.9	716	2	Q9H075	Q9h075	homo sapien
29	34	82.9	798	2	Q22763	Q22763	caenorhabdi
30	34	82.9	803	2	Q8TDW0	Q8tdw0	homo sapien
31	34	82.9	803	2	Q8R0N7	Q8r0n7	mus musculu
32	34	82.9	803	2	Q8R502	Q8r502	mus musculu
33	34	82.9	803	2	Q8C296	Q8c296	mus musculu
34	34	82.9	5149	2	Q9I179	Q9i179	pseudomonas
35	33	80.5	76	2	Q9KQC2	Q9kqc2	vibrio chol
36	33	80.5	242	2	Q8D5T9	Q8d5t9	vibrio vuln
37	33	80.5	247	2	Q7MCV7	Q7mcv7	vibrio vuln
38	33	80.5	260	2	Q9FMC8	Q9fmc8	arabidopsis

39	33	80.5	285	2	Q661W7	Q661w7 borrelia ga
40	33	80.5	324	1	YLPB_CAEEL	P34381 caenorhabdi
41	33	80.5	352	2	Q9XGZ6	Q9xgz6 arabidopsis
42	33	80.5	410	2	Q7R1K3	Q7r1k3 giardia lam
43	33	80.5	425	2	Q7ND52	Q7nd52 gloeobacter
44	33	80.5	465	2	Q8YDG4	Q8ydg4 brucella me
45	33	80.5	512	2	Q8UWD1	Q8uwd1 tetraodon n

# ALIGNMENTS

## RESULT 1

Q20231

ID Q20231 PRELIMINARY; PRT; 229 AA.  
AC Q20231;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein F40F8.3.  
GN ORFNames=F40F8.3;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA MacDougall R.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; Z69302; CAA93265.1; -.  
DR PIR; T22037; T22037.  
DR WormBase; WBGene00009576; F40F8.3.  
DR WormPep; F40F8.3; CE05843.  
KW Hypothetical protein.  
SQ SEQUENCE 229 AA; 24998 MW; 0447A8985EC1A4D3 CRC64;

Query Match 92.7%; Score 38; DB 2; Length 229;  
Best Local Similarity 77.8%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
:|||||:  
Db 209 ILSGLSDTI 217

## RESULT 2

Q8NYJ2

ID Q8NYJ2 PRELIMINARY; PRT; 394 AA.

AC Q8NYJ2;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE MW0207 protein.  
 GN OrderedLocusNames=MW0207;  
 OS Staphylococcus aureus (strain MW2).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MW2;  
 RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;  
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
 RA Yamamoto K., Hiramatsu K.;  
 RT "Genome and virulence determinants of high virulence community-  
 RT acquired MRSA.";  
 RL Lancet 359:1819-1827(2002).  
 CC -!- SIMILARITY: Belongs to the thiolase family.  
 DR EMBL; AP004822; BAB94072.1; -.  
 DR HSSP; P27796; 1AFW.  
 DR InterPro; IPR002155; Thiolase.  
 DR Pfam; PF02803; Thiolase\_C; 1.  
 DR Pfam; PF00108; Thiolase\_N; 1.  
 DR TIGRFAMs; TIGR01930; AcCoA-C-Actrans; 1.  
 DR PROSITE; PS00098; THIOLASE\_1; 1.  
 DR PROSITE; PS00737; THIOLASE\_2; 1.  
 DR PROSITE; PS00099; THIOLASE\_3; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 394 AA; 41807 MW; 25E7089FB91499A2 CRC64;

Query Match 92.7%; Score 38; DB 2; Length 394;  
 Best Local Similarity 88.9%; Pred. No. 24;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 ||:|||||  
 Db 75 LLTGLSDTV 83

#### RESULT 3

##### Q6GCP4

ID Q6GCP4 PRELIMINARY; PRT; 394 AA.  
 AC Q6GCP4;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Putative thiolase.  
 GN OrderedLocusNames=SAS0207;  
 OS Staphylococcus aureus (strain MSSA476).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=282459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15213324; DOI=10.1073/pnas.0402521101;  
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,

RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,  
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,  
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,  
 RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,  
 RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K.M., Sanders M.,  
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,  
 RA Spratt B.G., Parkhill J.;  
 RT "Complete genomes of two clinical *Staphylococcus aureus* strains:  
 RT evidence for the rapid evolution of virulence and drug resistance.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
 CC -!- SIMILARITY: Belongs to the thiolase family.  
 DR EMBL; BX571857; CAG41975.1; -.  
 DR InterPro; IPR002155; Thiolase.  
 DR Pfam; PF02803; Thiolase\_C; 1.  
 DR Pfam; PF00108; Thiolase\_N; 1.  
 DR TIGRFAMs; TIGR01930; AcCoA-C-AcTrans; 1.  
 DR PROSITE; PS00098; THIOLASE\_1; 1.  
 DR PROSITE; PS00737; THIOLASE\_2; 1.  
 DR PROSITE; PS00099; THIOLASE\_3; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 394 AA; 41807 MW; 25E7089FB91499A2 CRC64;

Query Match 92.7%; Score 39; DB 2; Length 394;  
 Best Local Similarity 88.9%; Pred. No. 24;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 ||:|||||  
 Db 75 LLTGLSDTV 83

#### RESULT 4

Q7RHZ2

ID Q7RHZ2 PRELIMINARY; PRT; 635 AA.  
 AC Q7RHZ2;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=PY03841;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=17XNL;  
 RX PubMed=12368865; DOI=10.1038/nature01099;  
 RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,  
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,  
 RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,  
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,  
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,  
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,

RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria  
 RT parasite *Plasmodium yoelii yoelii*.";  
 RL Nature 419:512-519(2002).  
 CC -!- SIMILARITY: Belongs to the DEAD box helicase family.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABL01001138; EAA15614.1; -.  
 DR HSSP; Q58083; 1HV8.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; Helicase\_C; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 KW ATP-binding; Helicase; Hydrolase; Hypothetical protein.  
 SQ SEQUENCE 635 AA; 73509 MW; DDFFF20CC5A38937 CRC64;

Query Match. 92.7%; Score 38; DB 2; Length 635;  
 Best Local Similarity 77.8%; Pred. No. 39;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 40;

QY 1 LLSGLSDTV 9  
 |:|||||:  
 Db 579 LISGLSDTI 587

# RESULT 5

Q93931

ID Q93931 PRELIMINARY; PRT; 270 AA.  
 AC Q93931;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein R07H5.3a.  
 GN Name=R07H5.3a; ORFNames=R07H5.3;  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C.elegans*: A platform for  
 RT investigating biology."; *Science* 282:2012-2018(1998).  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;

RA Lennard N.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; Z81107; CAB03226.2; -.  
 DR PIR; T24036; T24036.  
 DR WormBase; WBGene00011123; R07H5.3.  
 DR WormPep; R07H5.3a; CE35186.  
 DR InterPro; IPR006729; DUF598.  
 DR InterPro; IPR009024; MCR\_fer\_like.  
 DR Pfam; PF04635; DUF598; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 270 AA; 29981 MW; BFE2B2B3497613FA CRC64;

Query Match 90.2%; Score 37; DB 2; Length 270;  
 Best Local Similarity 77.8%; Pred. No. 26;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
 ||:|||||:  
 Db 213 LLTGLSDTI 221

# RESULT 6

## Q7JL05

ID Q7JL05 PRELIMINARY; PRT; 295 AA.  
 AC Q7JL05;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 ET 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein R07H5.3b.  
 GN Name=R07H5.3b; ORFNames=R07H5.3;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Lennard N.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; Z81107; CAE45743.1; -.  
 DR WormBase; WBGene00011123; R07H5.3.  
 DR WormPep; R07H5.3b; CE12554.  
 DR InterPro; IPR006729; DUF598.  
 DR InterPro; IPR009024; MCR\_fer\_like.  
 DR Pfam; PF04635; DUF598; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 295 AA; 32924 MW; E65D262A9B067419 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 295;



Best Local Similarity 77.8%; Pred. No. 29;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
||:|||||:  
Db 238 LLTGLSDTI 246

RESULT 7

Q9C8F3

ID Q9C8F3 PRELIMINARY; PRT; 75 AA.  
AC Q9C8F3;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein T15P17.13.  
GN Name=T15P17.13;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,  
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Town C.D., Kaul S.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AC025782; AAG51257.1; -.  
DR PIR; A86487; A86487.  
KW Hypothetical protein.  
SQ SEQUENCE 75 AA; 8254 MW; 4664A26082F1B106 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 75;  
Best Local Similarity 88.9%; Pred. No. 12;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
|||||:|  
Db 15 LLSGLSNTV 23

RESULT 8

Q9XEP0

ID Q9XEP0 PRELIMINARY; PRT; 548 AA.  
AC Q9XEP0;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Gibberellin MYB transcription factor.  
GN Name=GAMyb;  
OS Lolium temulentum (Darnel ryegrass).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Poaceae; Lolium.  
 OX NCBI\_TaxID=34176;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214457; PubMed=10198085; DOI=10.1104/pp.119.4.1271;  
 RA Gocal G.F., Poole A.T., Gubler F., Watts R.J., Blundell C., King R.W.;  
 RT "Long-day up-regulation of a GAMYB gene during Lolium temulentum  
 RT inflorescence formation.";  
 RL Plant Physiol. 119:1271-1278(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gocal G.F.W., Gubler F., Watts R.J., Blundell C., King R.W.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- SIMILARITY: Contains 2 Myb-like domains.  
 DR EMBL; AF114162; AAD31395.1; -.  
 DR HSSP; P06876; 1MBK.  
 DR TRANSFAC; T02856; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR009057; Homeodomain\_like.  
 DR InterPro; IPR001005; Myb\_DNA\_binding.  
 DR Pfam; PF00249; Myb\_DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW Nuclear protein.  
 SQ SEQUENCE 548 AA; 59303 MW; 6ECC1D188C5460A9 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 548;  
 Best Local Similarity 88.9%; Pred. No. 89;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 || |||||  
 Db 244 LLPGLSDTV 252

# RESULT 9

Q8NU44

ID Q8NU44 PRELIMINARY; PRT; 122 AA.  
 AC Q8NU44; Q6M8P9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Predicted membrane protein (PUTATIVE INNER MEMBRANE PROTEIN).  
 GN OrderedLocusNames=Cgl0103, cg0135;  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
 RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;  
 RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,  
 RA Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,  
 RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,  
 RA McHardy A.C., Meyer F., Moeckel B., Pfefferle W., Puehler A.,  
 RA Rey D.A., Rueckert C., Rupp O., Sahm H., Wendisch V.F., Wiegraebe I.,  
 RA Tauch A.;  
 RT "The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence  
 RT and its impact on the production of L-aspartate-derived amino acids  
 RT and vitamins.";  
 RL J. Biotechnol. 104:5-25(2003).  
 DR EMBL; AP005274; BAB97496.1; -.  
 DR EMBL; BX927148; CAF18671.1; -.  
 DR InterPro; IPR003807; DUF202.  
 DR Pfam; PF02656; DUF202; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 122 AA; 13246 MW; B715A4DD87F971CA CRC64;

Query Match 85.4%; Score 35; DB 2; Length 122;  
 Best Local Similarity 87.5%; Pred. No. 31;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSGLSDTV 9  
 :|||||  
 Db 52 ISGLSDTV 59

#### RESULT 10

##### Q8VFH5

ID Q8VFH5 PRELIMINARY; PRT; 312 AA.  
 AC Q8VFH5;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Olfactory receptor MOR111-6.  
 GN Name=Olfr775;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21676863; PubMed=11802173;  
 RA Zhang X., Firestein S.;  
 RT "The olfactory receptor gene superfamily of the mouse.";  
 RL Nat. Neurosci. 5:124-133(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21864068; PubMed=11875048; DOI=10.1093/hmg/11.5.535;  
 RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tonnes-Priddy L.,  
 RA Trask B.J.;  
 RT "Different evolutionary processes shaped the mouse and human olfactory  
 RT receptor gene families.";

RL Hum. Mol. Genet. 11:535-546(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 DR EMBL; AY073551; AAL61214.1; -.  
 DR MGD; MGI:3030609; Olfr775.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004984; F:olfactory receptor activity; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PRINTS; PR00245; OLFACTORYR.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Receptor; Transmembrane.  
 SQ SEQUENCE 312 AA; 35498 MW; B33931040DA1498D CRC64;

Query Match 85.4%; Score 35; DB 2; Length 312;  
 Best Local Similarity 87.5%; Pred. No. 81;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLSGLSDT 8  
 ||||:||||  
 Db 11 LLSGISDT 18

# RESULT 11

## Q7TRI4

ID Q7TRI4 PRELIMINARY; PRT; 312 AA.  
 AC Q7TRI4;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Olfactory receptor Olfr775.  
 GN Name=Olfr775;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22974002; PubMed=14611657; DOI=10.1186/gb-2003-4-11-r71;  
 RA Young J.M., Shykind B.M., Lane R.P., Tonnes-Priddy L., Ross J.A.,  
 RA Walker M., Williams E.M., Trask B.J.;  
 RT "Odorant receptor expressed sequence tags demonstrate olfactory  
 RT expression of over 400 genes, extensive alternate splicing and unequal  
 RT expression levels."  
 RL Genome Biol. 4:R71-R71(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Sanders K.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 DR EMBL; AY317950; AAP71272.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Receptor; Transmembrane.  
 SQ SEQUENCE 312 AA; 35464 MW; 193936B2DB7A243B CRC64;

Query Match 85.4%; Score 35; DB 2; Length 312;  
 Best Local Similarity 87.5%; Pred. No. 81;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLSGLSDT 8  
 ||||:|||  
 Db 11 LLSGISDT 18

# RESULT 12

Q71ZR1

ID Q71ZR1 PRELIMINARY; PRT; 391 AA.  
 AC Q71ZR1;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Major facilitator family transporter.  
 GN OrderedLocusNames=LMOF2365\_1428;  
 OS Listeria monocytogenes (serotype 4b / strain F2365).  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=265669;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15115801; DOI=10.1093/nar/gkh562;  
 RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,  
 RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,  
 RA Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,  
 RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,  
 RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,  
 RA Forberger H.A., Tran B., Kathariou S., Wonderling L.D., Uhlich G.A.,  
 RA Bayles D.O., Luchansky J.B., Fraser C.M.;  
 RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-  
 RT borne pathogen Listeria monocytogenes reveal new insights into the  
 RT core genome components of this species."  
 RL Nucleic Acids Res. 32:2386-2395(2004).  
 DR EMBL; AE017326; AAT04203.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0015904; P:tetracycline transport; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR007114; MFS.  
 DR InterPro; IPR001958; TCR\_TetA.  
 DR PRINTS; PR01035; TCRTETA.

DR PROSITE; PS50850; MFS; 1.  
KW Complete proteome.  
SQ SEQUENCE 391 AA; 42049 MW; 0A5D5BA43B140F13 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 391;  
Best Local Similarity 77.8%; Pred. No. 1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 83 LLCGLSDTI 91

RESULT 13

Q9F427

ID Q9F427 PRELIMINARY; PRT; 392 AA.  
AC Q9F427;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Multidrug efflux transporter.  
GN Name=mdet11;  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LO28;  
RA Perez-Diaz J.C.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ012115; CAC13963.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0015904; P:tetracycline transport; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR001958; TCR\_TetA.  
DR PRINTS; PR01035; TCR\_TetA.  
DR PROSITE; PS50850; MFS; 1.  
SQ SEQUENCE 392 AA; 42811 MW; E59A575EDDCCCC33 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 392;  
Best Local Similarity 77.8%; Pred. No. 1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLSGLSDTV 9  
|| |||||:  
Db 83 LLCGLSDTI 91

RESULT 14

Q92BU8

ID Q92BU8 PRELIMINARY; PRT; 397 AA.  
AC Q92BU8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Lin1446 protein.  
 GN OrderedLocusNames=lin1446;  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of Listeria species."  
 RL Science 294:849-852(2001).  
 DR EMBL; AL596168; CAC96677.1; -.  
 DR PIR; AE1613; AE1613.  
 DR ListiList; LIN1446; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0015904; P:tetracycline transport; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR007114; MFS.  
 DR InterPro; IPR001958; TCR\_TetA.  
 DR PRINTS; PR01035; TCRTETA.  
 DR PROSITE; PS50850; MFS; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 397 AA; 42777 MW; E78B9B959E6F9FAA CRC64;

Query Match 85.4%; Score 35; DB 2; Length 397;  
 Best Local Similarity 77.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 || |||||:  
 Db 89 LLCGLSDTI 97

# RESULT 15

Q9X769

ID Q9X769 PRELIMINARY; PRT; 397 AA.  
 AC Q9X769;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Multidrug resistance transporter (Lmol409 protein).  
 GN Name=lltB; OrderedLocusNames=lmol409;  
 OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99271176; PubMed=10339818;  
 RA Huillet E.E.H., Larpin S., Pardon P., Berche P.;  
 RT "Identification of a new locus in *Listeria monocytogenes* involved in  
 RT cellobiose-dependent repression of hly expression.";  
 RL FEMS Microbiol. Lett. 174:265-272(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of *Listeria* species.";   
 RL Science 294:849-852(2001).  
 DR EMBL; AJ009627; CAB43715.1; -.  
 DR EMBL; AL591979; CAC99487.1; -.  
 DR PIR; AI1250; AI1250.  
 DR PIR; T46710; T46710.  
 DR ListiList; LMO1409; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0015904; P:tetracycline transport; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR007114; MFS.  
 DR InterPro; IPR001958; TCR\_TetA.  
 DR PRINTS; PR01035; TCRTETA.  
 DR PROSITE; PS50850; MFS; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 397 AA; 42726 MW; 22690B5C8D6F45EA CRC64;

Query Match 85.4%; Score 35; DB 2; Length 397;  
 Best Local Similarity 77.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLSGLSDTV 9  
 || |||||:  
 Db 89 LLCGLSDTI 97

Search completed: February 10, 2005, 15:57:41  
 Job time : 67.662 secs